GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 22:26:00 ; Search time 44.2 Seconds
                                          (without alignments)
                                          88.680 Million cell updates/sec

Title: US-09-104-340-4
Perfect score: 1466
Sequence: 1 MDCQLSILLLSCSVLDSFG...........VPIGKCSCNAGYEERGFMCQ 271

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 143561 seqs, 14463640 residues

Total number of hits satisfying chosen parameters: 143561

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
                 4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	1466	100.0	983	1	US-08-167-919A-10	Sequence 10, Appl
2	1466	100.0	983	2	US-08-449-645A-21	Sequence 21, Appl
3	1466	100.0	983	2	US-08-702-367A-21	Sequence 21, Appl
4	1466	100.0	983	3	US-08-715-106-10	Sequence 10, Appl
5	1466	100.0	983	4	PCT-US95-04681-21	Sequence 21, Appl
6	1306	89.1	983	2	US-08-162-809-16	Sequence 16, Appl
7	1302	88.8	982	1	US-08-673-789-4	Sequence 4, Appl
8	1005	68.6	986	2	US-08-673-789-3	Sequence 3, Appl
9	1003	68.4	948	2	US-08-469-537A-101	Sequence 101, App
10	995	67.9	986	2	US-08-449-645A-15	Sequence 15, Appl
11	995	67.9	986	2	US-08-702-367A-15	Sequence 15, Appl
12	995	67.9	986	4	PCT-US95-04681-15	Sequence 15, Appl
13	995	67.9	1104	1	US-08-222-616-36	Sequence 36, Appl
14	995	67.9	1104	4	PCT-US95-04681-36	Sequence 36, Appl
15	990	67.5	877	2	US-08-673-789-2	Sequence 2, Appl
16	988	67.4	1005	2	US-08-469-537A-103	Sequence 103, App
17	987	67.3	928	1	US-08-442-248-2	Sequence 2, Appl
18	987	67.3	928	2	US-08-440-815-2	Sequence 2, Appl
19	984	67.1	991	2	US-08-449-645A-13	Sequence 13, Appl
20	984	67.1	991	2	US-08-702-367A-13	Sequence 13, Appl
21	984	67.1	991	4	PCT-US95-04681-13	Sequence 13, Appl
22	976.5	66.6	967	2	US-08-449-645A-30	Sequence 30, Appl
23	976.5	66.6	967	2	US-08-702-367A-30	Sequence 30, Appl
24	927.5	63.3	998	2	US-08-449-645A-17	Sequence 17, Appl
25	927.5	63.3	998	2	US-08-702-367A-17	Sequence 17, Appl
26	927.5	63.3	998	4	PCT-US95-04681-17	Sequence 17, Appl
27	926.5	63.2	610	4	PCT-US96-00419-3	Sequence 3, Appl
28	926.5	63.2	626	4	PCT-US96-00419-5	Sequence 5, Appl
29	926.5	63.2	998	4	PCT-US96-00419-2	Sequence 2, Appl
30	774	52.8	995	2	US-08-673-789-5	Sequence 5, Appl
31	772	52.7	995	2	US-08-162-809-18	Sequence 18, Appl
32	772	52.7	1011	2	US-08-162-809-12	Sequence 12, Appl
33	767	52.3	970	1	US-08-449-645A-11	Sequence 11, Appl
34	767	52.3	970	2	US-08-702-367A-11	Sequence 11, Appl
35	767	52.3	970	4	PCT-US95-04681-11	Sequence 11, Appl
36	763.5	52.1	984	2	US-08-162-809-11	Sequence 11, Appl
37	692	47.2	998	2	US-08-673-789-6	Sequence 6, Appl
38	692	47.2	998	2	US-08-449-645A-20	Sequence 20, Appl
39	692	47.2	998	2	US-08-702-367A-20	Sequence 20, Appl
40	681.5	46.5	710	4	PCT-US95-04681-20	Sequence 20, Appl
41	681.5	46.5	722	1	US-08-162-809-22	Sequence 22, Appl
42	681.5	46.5	744	2	US-08-162-809-4	Sequence 4, Appl
43	679	46.3	973	1	US-08-162-809-20	Sequence 20, Appl
44	679	46.3	988	1	US-08-162-809-14	Sequence 14, Appl
45	679	46.3	993	1	US-08-348-143-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-167-919A-10
; Sequence 10, Application US/08167919A
; Patent No. 5674691
; GENERAL INFORMATION:
; APPLICANT: Boyd, Andrew W.
; APPLICANT: Simpson, Richard J.
; APPLICANT: Wicks, Ian
; APPLICANT: Ward, Larry D.
; APPLICANT: Wilkinson, David
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,919A
; FILING DATE: 18-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK6841 (AU)
; FILING DATE: 21-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK9992 (AU)
; FILING DATE: 12-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00294
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Diglglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9159
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-167-919A-10

```
Query Match        100.0%;  Score 1466;  DB 1;  Length 983;
Best Local Similarity    100.0%;  Pred. No. 5.2e-146;
Matches 271;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60

Qy   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120

Qy  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180

Qy  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
         |||||||||||||||||||||||||||||||
Db  241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
```

RESULT 2
US-08-449-645A-21
Sequence 21, Application US/08449645A
Patent No. 5981245

GENERAL INFORMATION:

APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehaviland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-449-645A-21

```
Query Match        100.0%;  Score 1466;  DB 2;  Length 983;
Best Local Similarity    100.0%;  Pred. No. 5.2e-146;
Matches 271;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60

Qy   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120

Qy  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180

Qy  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
         |||||||||||||||||||||||||||||||
Db  241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
```

RESULT 3
US-08-702-367A-21
Sequence 21, Application US/08702367A
Patent No. 5981246

GENERAL INFORMATION:

APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehaviland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-367A-21

```
Query Match        100.0%;  Score 1466;  DB 2;  Length 983;
Best Local Similarity    100.0%;  Pred. No. 5.2e-146;
Matches 271;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE   60

Qy   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120

Qy  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY   180

Qy  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240
```

```
Qy  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240

Qy  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
        |||||||||||||||||||||||||||||||
Db  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 4
US-08-715-106-10
Sequence 10, Application US/08715106
Patent No. 6020306
GENERAL INFORMATION:
  APPLICANT: Boyd, Andrew W.
  APPLICANT: Simpson, Richard J.
  APPLICANT: Wicks, Ian
  APPLICANT: Ward, Larry D.
  APPLICANT: Wilkinson, David
  TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
  TITLE OF INVENTION: AND USE THEREOF
  NUMBER OF SEQUENCES: 14
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Scully, Scott, Murphy & Presser
  STREET: 400 Garden City Plaza
  CITY: Garden City
  STATE: New York
  COUNTRY: U.S.A.
  ZIP: 11530
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/715,106
  FILING DATE: 18-SEP-1996
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/167,919
  FILING DATE: 18-APR-1994
  APPLICATION NUMBER: PK6841 (AU)
  FILING DATE: 21-JUN-1991
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: PK9992 (AU)
  FILING DATE: 12-DEC-1991
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: PCT/AU92/00294
  FILING DATE: 19-JUN-1992
  ATTORNEY/AGENT INFORMATION:
  NAME: DiGiglio, Frank S.
  REGISTRATION NUMBER: 31,346
  REFERENCE/DOCKET NUMBER: 9159
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 742-4343
  TELEFAX: (516) 742-4366
  TELEX: 230 901 SANS UR
  SEQUENCE CHARACTERISTICS:
  INFORMATION FOR SEQ ID NO: 10:
  LENGTH: 983 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-715-106-10

Query Match   100.0%;   Score 1466; DB 3; Length 983;
Best Local Similarity   100.0%;   Pred. No. 5.2e-146;
Matches 271; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60

Qy   61 HYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 HYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120

Qy  121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180

Qy  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240

Qy  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
        |||||||||||||||||||||||||||||||
Db  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 5
PCT-US95-04681-21
Sequence 21, Application PC/TUS9504681
GENERAL INFORMATION:
  APPLICANT: Fox, Gary M.
  TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
  TITLE OF INVENTION: Kinases
  NUMBER OF SEQUENCES: 28
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Amgen Patent Operations/RBW
  STREET: 1840 Dehavilland Drive
  CITY: Thousand Oaks
  STATE: California
  COUNTRY: USA
  ZIP: 91320
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/04681
  FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
  NAME: Winter, Robert B.
  REFERENCE/DOCKET NUMBER: A-287
  INFORMATION FOR SEQ ID NO: 21:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 983 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: protein
PCT-US95-04681-21

Query Match   100.0%;   Score 1466; DB 4; Length 983;
Best Local Similarity   100.0%;   Pred. No. 5.2e-146;
Matches 271; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60

Qy   61 HYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 HYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120

Qy  121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180

Qy  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
```

```
Qy  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
        |||||||||||||||||||||||||||||||
Db  241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 6
US-08-162-809-16
; Sequence 16, Application US/0816809
; Patent No. 5457048
; GENERAL INFORMATION:
;   APPLICANT: Pasquale, Elena B.
;   APPLICANT: Sajjadi, Fereydoun G.
;   TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
;   TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
;   NUMBER OF SEQUENCES: 26
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: CAMPBELL AND FLORES
;   STREET: 4370 La Jolla Village Drive, Suite 700
;   CITY: San Diego
;   STATE: California
;   COUNTRY: United States of America
;   ZIP: 92122
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/162,809
;   FILING DATE:
;   CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Campbell, Cathryn A.
;   REGISTRATION NUMBER: 31,815
;   REFERENCE/DOCKET NUMBER: P-LJ 9503
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 535-9001
;   TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 983 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-162-809-16

Query Match   89.1%; Score 1306; DB 1; Length 983;
Best Local Similarity   91.3%; Pred. No. 4e-129;
  Matches 240; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

```
Qy   9 LILSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTY 68
       ||| | : ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   8 LILLCAALGSAGRLSARPGNEVNLLDSKTIQGELGWISYPSHGWEEISGVDENYTPIRTY 67

Qy  69 QVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 128
       | :|||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db  68 QESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 127

Qy 129 DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGA 188
       |||| |||||||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db 128 DDHLAKFREHQFTKIDTWAADESFTQMDLGDRILKLNTEVREVGPVSKKGFYLAFQDVGA 187

Qy 189 CVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
       ||||||||||||||||||||||||||||||||| ||||| | | |||||:|||:|||||
Db 188 CVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNHSKEEPPRKMYCSTEG 247

Qy 249 EWLVPIGKCSCNAGYEERGFMCQ 271
       ||||||||| |||||||||| ||
Db 248 EWLVPIGKCLCNAGYEERGFACQ 270
```

RESULT 7
US-08-673-789-4
; Sequence 4, Application US/08673789
; Patent No. 5814479
; GENERAL INFORMATION:
;   APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
;   APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
;   APPLICANT: GEORGE, F.
;   TITLE OF INVENTION: EHK RECEPTOR LIKE
;   TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
;   TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
;   TITLE OF INVENTION: METHODS
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN & FINNEGAN
;   STREET: 345 PARK AVENUE
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: USA
;   ZIP: 10154
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: FLOPPY DISK
;   COMPUTER: IBM PC COMPATIBLE
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WORDPERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/673,789
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/177,812
;   FILING DATE: 04-JAN-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: CAROL M. GRUPPI
;   REGISTRATION NUMBER: 37,341
;   REFERENCE/DOCKET NUMBER: 2026-4105
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 758-4800
;   TELEFAX: (212) 751-6849
;   TELEX: 421792
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 982
;   TYPE: AMINO ACID
;   STRANDEDNESS: UNKNOWN
;   TOPOLOGY: UNKNOWN
US-08-673-789-4

Query Match   88.8%; Score 1302; DB 2; Length 982;
Best Local Similarity   91.3%; Pred. No. 1e-128;
  Matches 240; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

```
Qy   9 LILSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTY 68
       ||| | : ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   8 LILLCAALGSAGRLSARPGNEVNLLDSKTIQGELGWISYPSHGWEEISGVDENYTPIRTY 67

Qy  69 QVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 128
       | :|||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db  68 QESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 127

Qy 129 DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGA 188
       |||| |||||||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db 128 DDHLAKFREHQFTKIDTWAADESFTQMDLGDRILKLNTEVREVGPVSKKGFYLAFQDVGA 187

Qy 189 CVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
       ||||||||||||||||||||||||||||||||| ||||| | | |||||:|||:|||||
Db 188 CVALVSVRYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNHSKEEPPRKMYCSTEG 247

Qy 249 EWLVPIGKCSCNAGYEERGFMCQ 271
       ||||||||| |||||||||| ||
Db 248 EWLVPIGKCLCNAGYEERGFACQ 270
```

RESULT 8
US-08-673-789-3
; Sequence 3, Application US/0867378789
; Patent No. 5814479
; GENERAL INFORMATION:
; APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
; APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
; APPLICANT: GEORGE, F.
; TITLE OF INVENTION: BSK RECEPTOR LIKE
; TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,789
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/177,812
; FILING DATE: 04-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
US-08-673-789-3

Query Match   68.6%; Score 1005; DB 2; Length 986;
Best Local Similarity 73.0%; Pred. No. 2.2e-97;
Matches 181; Conservative 33; Mismatches 32; Indels 2; Gaps 2;

Qy   26 PSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
         |:|||||||||||||||:||||||| |||||||||||||||||||||||:||:|||||
Db   27 PANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRT 86

Qy   85 NMVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKEIFNLYYMESDDHGVKFREHQFTKID 144
         : ::| |||||||||||||||||||||||||||||||||||:||||||:| |
Db   87 DWITREGAQRVYIEIKFTLRDCNSLPGVMGTCKETFNLYYESDNDKERFIRESQFGKTD 146

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKRGFYLAFQDVGACVALVSVRVYFKKCPF 204
         ||||||:|||||||||||||||||:||:|||||||||||||||||||||||||||:||||
Db   147 TIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPL 206

Qy   205 TVRNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGY 263
         |:|||||||||| |||||||||||||||:||||||||||:|||||||||||||||||||
Db   207 TVRNLAQFPDIITGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGY 266

Qy   264 EERGFMCQ 271
         ||||:|| ||
Db   267 EERGFACQ 274

_____

RESULT 9
US-08-469-537A-101
; Sequence 101, Application US/0846953737A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisonpierre, et al.
; TITLE OF INVENTION: EHK AND ROR TYROSINE
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 948 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-469-537A-101

Query Match   68.4%; Score 1003; DB 2; Length 948;
Best Local Similarity 65.7%; Pred. No. 3.4e-97;
Matches 180; Conservative 39; Mismatches 37; Indels 18; Gaps 3;

Qy   3 COLSILILLSCSVLDSFGELIP----------QPSNEVNLLDSKTIQGELGWISYPSHGW 52
         |:| |:|:|||||||||||||||||:|||||:||
Db   4 CEVREFLL------QFGFFLPLLTAWTGDCSHVSNQVVLLDTSTVNGELGWITYPLNGW 56

Qy   53 EEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNMVPRNSAQKIYVELKFTLRDCNSIPLV 112
         ||||||||||||||||:|||||||||||||||:|| ::|||||||||||||||||||||
Db   57 DAITEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIYVEMKFTLRDCNSIPWV 116

Qy   113 LGTCKEIFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVG 172
         |||||||:|||||:|||||||||||||||||:| ::|||||||||||||||||||||||
Db   117 LGTCKEIFTLYYIESDESHGIKFKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREVG 176

Qy   173 PVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCV 231
         |: :|||||||||::||||||||||||:|| ||||||||||||||| ||||||||||||||
Db   177 PIERKGFYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCV 236

Qy   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 265
         ||:||||||||||||||||||||||||||||||
Db   267 ...

Db   237 KSSEERDTPKLYCGADGDWLYPLGRCICTTGYEE 270
           :|:|:|| ::|||:|:|:|| |:|| | |||

RESULT 10
US-08-449-645A-15
Sequence 15, Application US/0849645A
Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-449-645A-15

Query Match   67.9%; Score 995; DB 2; Length 986;
Best Local Similarity   68.8%; Pred. No. 2.5e-96;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
           | || | :| |||||::||||||||||||||| ||| : || |||||:|| |:|||
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
           ||||||| || |:::|| :|:|: :|| :|:|:||||||||||| ||:|:||||||:| ||
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTCKETFNLYYE 128

Qy   127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
           |:|:|| |:|:||| |||||||||||||||||||:|||||| ||| ::||||||||||||
Db   129 SDNDKERFIRENQFVKIDTIAADESFIQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV 188

Qy   187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEDPPRMYCS 245
           ||| |||:||||:||||||| ::|||||||||||||:|||:|||||:|:|:| |||||||
Db   189 GACIALVSVRVFYKKCPLTVRNLAQFPDTIGADTSSLVEVRGSCVNNSEEKDVPKMYCG 248

Qy   246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
           :|||||:|||||:||||||:|:|||||
Db   249 ADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 11
US-08-702-367A-15
Sequence 15, Application US/0870236A
Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine

; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-367A-15

Query Match   67.9%; Score 995; DB 2; Length 986;
Best Local Similarity   68.8%; Pred. No. 2.5e-96;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
           | || | :| |||||::||||||||||||||| ||| : || |||||:|| |:|||
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
           ||||||| || |:::|| :|:|: :|| :|:|:||||||||||| ||:|:||||||:| ||
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTCKETFNLYYE 128

Qy   127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
           |:|:|| |:|:||| |||||||||||||||||||:|||||| ||| ::||||||||||||
Db   129 SDNDKERFIRENQFVKIDTIAADESFIQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV 188

Qy   187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEDPPRMYCS 245
           ||| |||:||||:||||||| ::|||||||||||||:|||:|||||:|:|:| |||||||
Db   189 GACIALVSVRVFYKKCPLTVRNLAQFPDTIGADTSSLVEVRGSCVNNSEEKDVPKMYCG 248

Qy   246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
           :|||||:|||||:||||||:|:|||||
Db   249 ADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 12
PCT-US95-04681-15
Sequence 15, Application PCT/US9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04681-15

Query Match   67.9%;   Score 995; DB 4;   Length 986;
Best Local Similarity   68.8%; Pred. No. 2.5e-96;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVVIEIKFTLRDCNSLPGVWGTCKETFNLYYE 128

Qy 127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
Db 129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV 188

Qy 187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS 245
Db 189 GACIALVSVRVFYKKCPLTVRNLAQPFDTITGADTSSLVEVRGSCVNNSEEKDVPRKMYCG 248

Qy 246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
Db 249 ADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 13
US-08-222-616-36
; Sequence 36, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935
; FILING DATE: 22-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 821P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-222-616-36

Query Match   67.9%;   Score 995; DB 1;   Length 1104;
Best Local Similarity   68.8%; Pred. No. 3e-96;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVVIEIKFTLRDCNSLPGVWGTCKETFNLYYE 128

Qy 127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
Db 129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV 188

Qy 187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS 245
Db 189 GACIALVSVRVFYKKCPLTVRNLAQPFDTITGADTSSLVEVRGSCVNNSEEKDVPRKMYCG 248

Qy 246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
Db 249 ADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 14
PCT-US95-04228-36
; Sequence 36, Application PC/TUS9504228
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1104 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-04228-36

Query Match 67.9%; Score 995; DB 4; Length 1104;
Best Local Similarity 68.8%; Pred. No. 3e-96;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

```
Qy  10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR  66
            |:| |  |||||||||||||||||||||||||||||| ||||||||||||||||
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR  68

Qy  67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETFNLYYME  126
           ||||||||| |||||||| |:|:| ||||||||||||||||||||||||||||||||||
Db  69 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTICKETFNLYYE  128

Qy 127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV  186
           |:|:|                     | |||||||||||||||||||||||||||||||
Db 129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV  188

Qy 187 GACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS  245
           ||||:||||||||||||||||:|||||||||| |||||||||||||||||||||||||
Db 189 GACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVRGSCVNNSEEKDVPKMYCG  248

Qy 246 TEGEWLVPIGKCSCNAGYEERGFMCQ  271
            :|||||||||| | ||||||||||
Db 249 ADGEWLVPIGNCLCNAGHEERSGECQ  274
```

APPLICATION NUMBER: US/08/673,789
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: 421792
TELEX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 877
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-2

Query Match 67.5%; Score 990; DB 2; Length 877;
Best Local Similarity 66.2%; Pred. No. 7.2e-96;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
Qy  10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ  69
           |:||  |||||||||||||||||||| |: ||||||| ||||||| |||||| | ||||
Db  46 LILCAALRT---LLASPSNEVNLLDSRIVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy  70 VCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETFNLYYMESDD  129
           |||||| ::|: |||| |||::||||||||||||||||:| | ||||||||| ||:||||
Db 103 VCKVMEQNQNNWLLITSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNNYYFESDD 162

Qy 130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC  189
            |: | |::|| |:||||||||||| |:|||||::|||||||:||||:||||||||||||
Db 163 ENGRSIKENQYIKDTIAADESFTELDLGDRVMKLNTEIRDVGPLSKKGFYLAFQDVGAC  222

Qy 190 VALVSVSRVYFKKCPFTVKNLAMFPDTVP--MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
           :||||| |||:||| :||:|| ||||||  |||| ||||| ||||:| ||||:|| ||:||
Db 223 IALVSVRVYYKKCPSVVRHLAIFPDTITGADSSOLLEVSGSCVNHSVIDDPPKMHCSAEG  282

Qy 249 EWLVPIGKCSCNAGYEERGFMCQ  271
           |||||||| | |||||| | ||
Db 283 EWLVPIGKCMCKRAGYEEKNGTCQ  305
```

US-08-673-789-2

RESULT 15
US-08-673-789-2
Sequence 2, Application US/0867789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

us-09-104-340-4.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Om protein - protein search, using sw model

Run on: May   9, 2000, 21:52:05 ; Search time 50.71 Seconds
                                              (without alignments)
                                              126.581 Million cell updates/sec

Title: US-09-104-340-4
Perfect score: 1466
Sequence: 1 MDCQLSTILLLSCSVLDSFG.........VPIGKCSCNAGYEERGFMCQ 271

Scoring table: BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   A_Geneseq_36:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	DB	ID	Description
1	1466	100.0	983	R31466	HEK polypeptide. R
2	1306	89.1	983	R75711	Eph-related PTK Ce
3	1003	68.4	948	W83148	Rat receptor tyros
4	995	67.9	986	R85936	Protein tyrosine-k
5	995	67.9	986	R85091	EPH-like receptor
6	990	67.5	877	W71628	Mouse Bsk receptor
7	988	67.4	1005	W83147	Rat receptor tyros
8	987	67.3	928	R97853	Rat REK7 eph-relat
9	984	67.1	991	R85090	EPH-like receptor
10	927.5	63.3	998	R85092	EPH-like receptor
11	926.5	63.2	610	W03422	Mouse developmenta
12	926.5	63.2	626	W03423	Mouse developmenta
13	926.5	63.2	998	W03421	Mouse developmenta
14	772	52.7	995	R75712	Eph-related PTK Ce
15	772	52.7	1011	R75709	Eph-related PTK Ce
16	772	52.7	970	R85089	EPH-like receptor
17	767	52.3	994	R87018	EPH-like receptor
18	767	52.3	994	W26366	Receptor tyrosine
19	761.5	51.9	984	R44513	Mouse Nuk tyrosine
20	692	47.2	990	R51899	elk. Expression of
21	681.5	46.5	710	R75714	Human embryonal ki
22	681.5	46.5	722	R75705	Eph-related PTK Ce
23	681.5	46.5	744	R75713	Eph-related PTK Ce
24	679	46.3	973	R75708	WPi. 93-036373/04.
25	679	46.3	988	R75710	Eph-related PTK Ce
26	679	46.3	993	R75843	Eph-related PTK Ce
27	679	46.3	993	R75844	Protein p140 cDNA
28	622	42.4	973	R75707	Eph-related PTK Ce
29	616	42.0	977	W19258	Embryonic stem cel
30	588	40.1	951	R75704	Eph-related CEK6.
31	569.5	38.8	849	R75706	Eph-related PTK Ce
32	528.5	36.1	1006	W72256	Human receptor typ
33	528.5	36.1	1006	W70525	Human thymus recep
34	528.5	36.1	1021	W70526	Human thymus recep

ALIGNMENTS

RESULT   1

ID   R31466 standard; Protein; 983 AA.
AC   R31466;
DT   24-MAY-1993 (first entry)
DE   HEK polypeptide.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; Hela; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FH
FT   peptide   1..20
FT   /note= "Signal peptide"
FT   protein   21..39
FT   /note= "Purified HEK protein #1"
FT   protein   840..860
FT   /note= "Purified HEK protein #2"
FT   modified_site   232..234
FT   /note= "N-link glycosylation"
FT   modified_site   337..339
FT   /note= "N-link glycosylation"
FT   modified_site   391..393
FT   /note= "N-link glycosylation"
FT   modified_site   404..406
FT   /note= "N-link glycosylation"
FT   modified_site   493..495
FT   /note= "N-link glycosylation"
FT   domain   542..565
FT   /note= "Transmembrane domain"
FT   binding_site   628
FT   /note= "ATP binding site"
FT   binding_site   630
FT   /note= "ATP binding site"
FT   binding_site   633
FT   /note= "ATP binding site"
FT   modified_site   779
FT   /note= "Putative autophosphorylation site"
PN   W09300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU0294.
PR   21-JUN-1991; AU-006841.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-EEK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence represets human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour Hela. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the

35	493	33.6	522	1	R76466	Non-differentiated
36	493	33.6	522	1	R94653	Receptor type tyro
37	493	33.6	522	1	W06330	Receptor-type tyro
38	493	33.6	522	1	W11303	Receptor-type tyro
39	493	33.6	972	1	R76468	Mature non-differe
40	493	33.6	972	1	W06331	Full-length recept
41	493	33.6	987	1	R85930	Protein tyrosine-k
42	493	33.6	987	1	R89263	Human non-differen
43	493	33.6	987	1	R94652	Receptor type tyro
44	493	33.6	987	1	W06335	Full length recept
45	493	33.6	987	1	W11304	Receptor-type tyro

```
CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence  983 AA;

Query Match             100.0%;  Score 1466;  DB 1;  Length 983;
Best Local Similarity   100.0%;  Pred. No. 6.5e-143;
Matches  271;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 MDCQLSILLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MDCQLSILLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60

Qy     61 HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETF 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETF 120

Qy    121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 NLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180

Qy    181 LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240

Qy    241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
          |||||||||||||||||||||||||||||||
Db    241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271


RESULT  2
ID   R75711 standard; Protein; 983 AA.
AC   R75711;
DT   11-NOV-1995  (first entry)
DE   Eph-related PTK Cek4.
KW   Cek4; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis.
OS   Gallus sp.
OS   WO951375-A.
PN   WO951375-A.
PD   08-JUN-1995.
PD   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB  Sajjadi FG;
WPI: 95-215256/28.
DR   N-PSDB; Q09059.
DR   WPI; 95-215256/28.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Disclosure; Page 85-89; 129pp; English.
CC   Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and
CC   embryonic tissues and also in the adult brain and retina.
SQ   Sequence  983 AA;

Query Match             89.1%;  Score 1306;  DB 1;  Length 983;
Best Local Similarity   91.3%;  Pred. No. 2.3e-126;
Matches  240;  Conservative   9;  Mismatches  14;  Indels   0;  Gaps   0;

Qy      9 LLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTY 68
          ||||:||||||| |:||||||||||||||||||||||||||||||||||||||||||||
Db      8 LLLCAALGSAGRLSARPGNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTY 67

Qy     69 QVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETFNLYYMESD 128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 QESNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 127

Qy    129 DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGA 188
          |||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db    128 DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGA 187

Qy    189 CVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db    188 CVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNHSKEEPPKMYCSTEG 247

Qy    249 EWLVPIGKCSCNAGYEERGFMCQ 271
          ||||||||| |||||||||| ||
Db    248 EWLVPIGKCLCNAGYEERGFACQ 270


RESULT  3
ID   W83148 standard; Protein; 948 AA.
AC   W83148;
DT   11-FEB-1999  (first entry)
DE   Rat receptor tyrosine kinase Ehk-2.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis.
OS   Rattus sp.
OS   US564374-A.
OS   US564374-A.
PD   01-DEC-1998.
PD   06-JUN-1998; 469537.
PF   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Maslakowski P, Yancopoulos GD;
DR   WPI; 99-044584/04.
DR   N-PSDB; V70208.
PT   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
CC   Example; Fig 21; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence represents rat Ehk-2.
SQ   Sequence  948 AA;

Query Match             68.4%;  Score 1003;  DB 1;  Length 948;
Best Local Similarity   65.7%;  Pred. No. 4.6e-95;
Matches  180;  Conservative  39;  Mismatches  37;  Indels  18;  Gaps   3;

Qy      3 COLSILLLLSCSVLDSFGELIP------------QPSNEVNLLDSKTIQGELGWISYPSHGW 52
          |:|           |:||             |:|||||||||||||:|||||||||||:|
Db      4 CEVRDFLL-------QFGFFLPLLTAWTGDCSHVSNQVVLLDTSTVMGELGWKTYPLNGW 56

Qy     53 EEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLV 112
          ||:|||||||||||:|||||||||||:|||||||||||||||:|||||||||||||||:
Db     57 DAITEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIYEMKFTLRDCNSIPWV 116

Qy    113 LGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVG 172
          ||||||||||||:||:||||||||||||||||||||||||||||||||||||||||||:
Db    117 LGTCKETFLYYIBSDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREVG 176

Qy    173 PVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCV 231
          ||||||||||||:||||:|||||||||||||||:|:|||:|||| ||||||||||||||
Db    177 PIERKGFYLAFQDIGDACIALVSVRVYFKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCV 236

Qy    232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 265
          |:|:|||:||||||||||||||:|::|||||||
Db    237 KSSEERDTPKLYCGADGDWLVPLGRCICTTGYEE 270


RESULT  4
ID   R85936 standard; Protein; 986 AA.
AC   R85936;
DT   14-FEB-1996  (first entry)
DE   Protein tyrosine-kinase bpTK7.
```

KW   Protein tyrosine-kinase; PTK; bpTK7; agonist; cell growth;
KW   differentiation.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   peptide   1..19
FT   /label= Sig_peptide
FT   domain   20..547
FT   /label= Sig_peptide
FT   domain   548..570
FT   /label= Extracellular_domain
FT   domain   571..986
FT   /label= Transmembrane_domain
FT   /label= Intracellular_tyrosine-kinase_domain
PN   WO9527061-A1.
PD   12-OCT-1995;
PF   04-APR-1995; US-222616.
PR   04-APR-1994; US-222616.
PA   (GETH) GENENTECH INC.
PI   Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
DR   Wood WI;
PI   WPI; 95-366160/47.
DR   N-PSDB; T03100.
PT   Agonist antibodies which activate specific protein tyrosine
PT   kinase(s) - also activate chimeric proteins of kinase extracellular
PT   domain and Ig constant domain, useful for studying, and therapeutic
PT   modulation of, cell growth and differentiation.
PS   Disclosure; Page 95-99; 125pp; English.
CC   DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC   to screen cDNA libraries to identify novel pTK genes. The bpTKs,
CC   bpTK1, bpTK2, bpTK3, bpTK4, bpTK5 and bpTK7 (R85924-28 and R85935,
CC   respectively) are expressed in human brain tissue and show homology
CC   to known pTKs. A full-length sequence for bpTK7 (R85936) was also
CC   obtd. bpTK7 can be used to design drugs that modulate pTK activity.
SQ   Sequence 986 AA;

Query Match   67.9%;   Score 995;   DB 1;   Length 986;
Best Local Similarity   68.8%;   Pred. No. 3.3e-94;
Matches 183; Conservative 37;   Mismatches 42;   Indels 4;   Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR   66
              :|: |||||||:|||||||||||||||||||||||| :|||||||||||||||||
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR   68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME   126
             |||||||||||||||||||:||:: | |:|||||||||||||||||||||||||||| |
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTCKETFNLYYE   128

Qy   127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV   186
             |:|: ||||||||||||||||||:|||:||:||||||||||||||:||||||||||||
Db   129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV   188

Qy   187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS   245
             |||::||:|||||||||:||||||:|||||| :||||||||||||||:||::|||||||
Db   189 GACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVERGSSCVNNSEEKDVPKMYCG   248

Qy   246 TEGEWLVPIGKCSCNAGYEERGFMCQ   271
             :||||||||| :||||||||:||||||
Db   249 ADGEWLVPIGNCLCNAGHEERSGECQ   274

PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-373799/48.
DR   N-PSDB; T02948.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc.; used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
CC   Claim 18; Page 62-65; 133pp; English.
CC   4 Novel human EPH-like protein tyrosine kinases, HEK5, HEK7,
CC   HEK8 and HEK11 (R85089-92), respectively, were identified following
CC   isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC   cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC   HEK receptors (or their soluble extracellular domains) are produced by
CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   assays), or to identify and purify ligands for HEK receptors, or
CC   therapeutically to modulate the activation of cell-associated
CC   receptors.
SQ   Sequence 986 AA;

Query Match   67.9%;   Score 995;   DB 1;   Length 986;
Best Local Similarity   68.8%;   Pred. No. 3.3e-94;
Matches 183; Conservative 37;   Mismatches 42;   Indels 4;   Gaps 3;

Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR   66
              :|: |||||||:|||||||||||||||||||||||| :|||||||||||||||||
Db   9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR   68

Qy   67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME   126
             |||||||||||||||||||:||:: | |:|||||||||||||||||||||||||||| |
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTCKETFNLYYE   128

Qy   127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV   186
             |:|: ||||||||||||||||||:|||:||:||||||||||||||:||||||||||||
Db   129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV   188

Qy   187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS   245
             |||::||:|||||||||:||||||:|||||| :||||||||||||||:||::|||||||
Db   189 GACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVERGSSCVNNSEEKDVPKMYCG   248

Qy   246 TEGEWLVPIGKCSCNAGYEERGFMCQ   271
             :||||||||| :||||||||:||||||
Db   249 ADGEWLVPIGNCLCNAGHEERSGECQ   274

RESULT 6
ID   W71628 standard; Protein; 877 AA.
AC   W71628;
PN   US5814479-A.
PD   29-SEP-1998;
PF   11-JUN-1996; 673789.
PR   04-JAN-1994; US-177812.
PR   11-JUN-1996; US-673789.
DE   Mouse Bsk receptor-like tyrosine kinase.
KW   Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW   neurodegenerative disease; limbic system neuron regeneration;
KW   chromosomal abnormality; degenerative growth; development disorder;
KW   viral infection; bacterial infection; Alzheimer's disease; epilepsy;
OS   Mus sp.
KW   schizophrenia; stroke; cerebral ischaemia.
PN   US5814479-A.
PF   11-JUN-1996; 673789.
PR   29-SEP-1998;
PR   04-JAN-1994; US-177812.
PR   11-JUN-1996; US-673789.
PA   (KROM/) KROMER L F.
PA   (KROM/) KROMER L F.
PA   (SCHU/) SCHULZ N T.
PA   (WOUD/) WOUDE G F V.
PA   (ZHOU/) ZHOU R.
PI   Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR   WPI; 98-541751/46.
DR   N-PSDB; V58192.

RESULT 5
ID   R85091 standard; Protein; 986 AA.
AC   R85091;
ID   R85091;
DT   16-APR-1996 (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK8.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK8;
KW   human eph-like kinase; therapy; diagnosis; vector; antibody.
OS   Homo sapiens.
PN   WO9528484-A1.
PD   26-OCT-1995;
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.

PT Isolated nucleic acid sequence encoding protein - used in Bsk
PT nucleic acid probes, used in detecting alterations in level of Bsk
PT messenger-RNA in biological samples isolated from mammal afflicted
PT with disease
PS Claim 1; Fig 2; 72pp; English.
CC The present sequence represents mouse Bsk, which is a receptor-like
CC tyrosine kinase. The nucleic acid sequence encoding Bsk can be used in
CC Bsk nucleic acid probes, which can be used in detecting alterations in
CC the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC from a mammal afflicted with a disease, such as neurodegenerative
CC diseases or disorders and neoplasms. The nucleic acid sequence can also
CC be delivered into the limbic system of patients with limbic system
CC neurodegenerative disease, disorder or injury, to promote or enhance
CC limbic system neuron regeneration or growth. Such neurodegenerative
CC diseases include, chromosomal abnormalities, degenerative growth and
CC development disorders, viral infections, bacterial infections, brain
CC injuries, neoplastic conditions, Alzheimer's disease, epilepsy,
CC schizophrenia, or stroke and cerebral ischaemia.
SQ Sequence 877 AA;

Query Match   67.5%; Score 990; DB 1; Length 877;
Best Local Similarity 66.2%; Pred. No. 9e-94;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
         |  ||  :  ||||| | :  ||||| : ||||: | |: |||||| ||:  || |||
Db   46 LLLCAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy   70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         |||:| :|| ||| :: :||||  :||||||||||||||||||| |||||||:||||||
Db  103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy  130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         :|| |||||:|||||||||||||| |||||| |||| |:|||||||||||||||||||
Db  163 ENGRSIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy  190 VALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         :||||| | |||| :||  ||||| | |||| ||:||||||||: :||: :|| ||
Db  223 IALVSVRVYYKKCPSVVRHLAIFPDTIIGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy  249 EWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||| ||||||| ||:||
Db  283 EWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 7
W83147
ID W83147 standard; Protein; 1005 AA.
AC W83147;
DT 11-FEB-1999 (first entry)
DE Rat receptor tyrosine kinase Ehk-1.
KW Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW binding protein; BDNF; NT-3; diagnosis.
OS Rattus sp.
PN US5843749-A.
PD 01-DEC-1998;
PF 06-JUN-1995; 469537.
PR 17-MAR-1995; US-406247.
PR 26-JUL-1991; US-735559.
PR 28-OCT-1993; US-144992.
PR 06-JUN-1995; US-469537.
PA (REGE-) REGENERON PHARM INC.
PI Maisonpierre PC, Masiakowski P, Yancopoulos GD;
DR WPI; 99-044584/04.
DR N-PSDB; V70207.
PT DNA encoding receptor tyrosine kinase proteins - and corresponding
PT proteins
PS Example; Fig 22; 194pp; English.
CC The present invention describes nucleic acid molecules for ror-1,
CC ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor

CC tyrosine kinases. The present sequence represents rat Ehk-1.
SQ Sequence 1005 AA;

Query Match   67.4%; Score 988; DB 1; Length 1005;
Best Local Similarity 66.2%; Pred. No. 1.8e-93;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
         |  ||  :  ||||| | :  ||||| : ||||: | |: |||||| ||:  || |||
Db   46 LLLCAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy   70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         |||:| :|| ||| :: :||||  :||||||||||||||||||| |||||||:||||||
Db  103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy  130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         :|| |||||:|||||||||||||| |||||| |||| |:|||||||||||||||||||
Db  163 ENGRIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy  190 VALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         :||||| | |||| :||  ||||| | |||| ||:||||||||: :||: :|| ||
Db  223 IALVSVRVYYKKCPSVVRHLAVFPDTIITGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy  249 EWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||| ||||||| ||:||
Db  283 EWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 8
R97853
ID R97853 standard; Protein; 928 AA.
AC R97853;
DT 05-JAN-1997 (first entry)
DE Rat REK7 eph-related tyrosine kinase receptor.
KW REK7; eph-related tyrosine kinase receptor; AL-1; ligand;
KW neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis.
OS Rattus sp.
FH Key   Location/Qualifiers
FT peptide   1..57
FT   /label= Sig_peptide
FT protein   58..928
FT   /label= Mat_protein
FT domain   58..462
FT   /label= Extracellular_domain
PN WO9613518-A1.
PD 09-MAY-1996;
PF 26-OCT-1995; U14016.
PR 26-OCT-1995; U14016.
PR 27-OCT-1994; US-330128.
PR 07-JUN-1995; US-486449.
PA (GETH) GENENTECH INC.
PI Caras IW, Winslow JW;
DR WPI; 96-239448/24.
DR N-PSDB; T18893.
PT AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
PT treatment and diagnosis of neuronal disorders and
PT angiogenesis-related conditions.
PS Example 1; Page 50-53; 75pp; English.
CC Rat REK7 (R97853) is an eph-related tyrosine kinase receptor, for
CC which AL-1 (see also W97854) is a ligand. Its amino acid sequence
CC was deduced from a cDNA clone (T18893) isolated from a hippocampal
CC cDNA library. An REK-IgG fusion was used to screen cultured cell
CC lines for surface expression of REK7-binding activity. Primers
CC (see also T44382-83) based on peptide sequences (R97856-59) of
CC isolated ligands were used to amplify human breast carcinoma BT20
CC cell cDNA, and an amplified fragment was used to screen a human
CC foetal brain cDNA library, leading to the isolation of AL-1 cDNA
CC (T18897).
SQ Sequence 928 AA;

Query Match   67.3%; Score 987; DB 1; Length 928;
Best Local Similarity 66.2%; Pred. No. 2e-93;

```
Matches 174;  Conservative 44;  Mismatches 41;  Indels 4;  Gaps 2;

Qy       10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
            |||||||||||||:|:|||||:|||||||||||||||||:||||||:|||||:||:||||
Db       46 LLICAALRT---LLASPSNEVNLLDSRTVLGDLGWIAFPKNGWEEIGEVDENYPPIHTYQ 102

Qy       70 VCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
            || :||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      103 VCKVMEQNQNNWLLISWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy      130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
            :|:  ||||||||||||||:|||||||:||||||||||||||||||||||||||||||||
Db      163 ENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKRGFYLAFQDVGAC 222

Qy      190 VALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
            ||||||||||||||||||||||||||||| ||:|:|||||||||||||:|||||||||||
Db      223 IALVSVRVYIKKCPSVVRHLAVFPDIITGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy      249 EWLVPIGKCSCNAGYEERGFMCQ 271
            ||||||||||||||||||||||| |
Db      283 EWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 9

```
ID   R85090 standard; Protein; 991 AA.
AC   R85090;
DT   16-APR-1996  (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK7.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK7;
KW   human eph-like kinase; therapy; diagnosis; vector; antibody.
OS   Homo sapiens.
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; U04681.
PR   14-APR-1995; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-373799/48.
DR   N-PSDB; T02947.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 18; Page 54-57; 133pp; English.
CC   4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC   HEK8 and HEK11 (R85089-92), respectively, were identified following
CC   isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC   cDNA library.  HEK5, HEK7 and HEK8 show extensive homology to the
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   HEK11 shows no homology to any known EPH-like receptor.  Recombinant
CC   HEK receptors (or their soluble extracellular domains) are produced by
CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   assays), or to identify and purify ligands for HEK receptors, or
CC   therapeutically to modulate the activation of cell-associated
CC   receptors.  Soluble HEK7 receptor may primarily affect
CC   proliferation and/or differentiation of brain cells.
SQ   Sequence 991 AA;

Query Match    67.1%;  Score 984;  DB 1;  Length 991;
Best Local Similarity  65.8%;  Pred. NO. 4.5e-93;
Matches 173;  Conservative  44;  Mismatches 42;  Indels 4;  Gaps 2;

Qy       10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
            |||||||||||||:|:|||||:||||||||||||||||||:|||||||:|||||:||:||
Db       20 LLICAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 76

Qy       70 VCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
            || :||||||||||||||||||||||||||||||||||||:|||||||||||:|||:||
Db       77 VCKVMEQONQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 136
```

RESULT 10

```
ID   R85092 standard; Protein; 998 AA.
AC   R85092;
DT   16-APR-1996  (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK11.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK11;
KW   human eph-like kinase; therapy; diagnosis; antibody; vector.
OS   Homo sapiens.
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-373799/48.
DR   N-PSDB; T02949.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 18; Page 71-75; 133pp; English.
CC   4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC   HEK8 and HEK11 (R85089-92), respectively, were identified following
CC   isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC   cDNA library.  HEK5, HEK7 and HEK8 show extensive homology to the
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   HEK11 shows no homology to any known EPH-like receptor.  Recombinant
CC   HEK receptors (or their soluble extracellular domains) are produced by
CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   assays), or to identify and purify ligands for HEK receptors, or
CC   therapeutically to modulate the activation of cell-associated
CC   receptors.
SQ   Sequence 998 AA;

Query Match    63.3%;  Score 927.5;  DB 1;  Length 998;
Best Local Similarity  68.4%;  Pred. No. 3.2e-87;
Matches 171;  Conservative  39;  Mismatches 37;  Indels 3;  Gaps 2;

Qy       25 QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
            :||:||||||||||||||||||:||||||||:||||||||||||||:|||||||||||||
Db       28 QAAKEVLLLDSKAQOTELEWISSPPNGWEEIGGLDENYTPIRTYQVCQVMEPNQONNWLRT 87

Qy       85 NWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKID 144
            || |||||||||||||||||||:|||||||||||||||||||:::|| ||||||||||
Db       88 NWISKGNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLYYFDYDTGRNIRENLYVKID 147

Qy      145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 204
            |||||||||||||||:|||||||||:| |||||||||||||||||
Db      148 TIAADESFTQGDLGERKMKLNTEVREIGPLSKKGFYLAFQDVGAC 207

Qy      205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE---EDPPRMYCSTEGEWLVPIGKCSCNA 261
            :|||||||||||| |||||||||||||:|||::|  |:||||||||||:||||||||:|
Db      208 IIENLAIFPDTVTGSEFSLVEVRGTCVSSAEEAENAPRMHCSAEGEWLVPIGKCICKA 267

Qy      262 GYEERGFMCQ 271
            ||:::|| |:
Db      257 EWLVPIGKCMCKAGYEEKNGTCQ 279
```

```
Qy      130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
            :|:  ||||||||||||||:|||||||:|||||||||||||||||||||||||||||||
Db      137 QNGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKRGFYLAFQDVGAC 196

Qy      190 VALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
            ||||||||||||||||||||||||||||| ||:|:|||||||||||||::||||||||||
Db      197 IALVSVRVYIKKCPSVVRHLAVFPDIITGADSSQLLEVSGSCVNHSVTDEPPKMHCSAEG 256

Qy      249 EWLVPIGKCSCNAGYEERGFMCQ 271
            ||||||||||||||||||||||| |
Db      257 EWLVPIGKCMCKAGYEEKNGTCQ 279
```

Left column — RESULT 11

```
Db   268  GYQQKGDTCE  277

RESULT 11
ID   W03422
AC   W03422;
DT   11-NOV-1996  (first entry)
DE   Mouse developmental kinase 1 MDK1 T1.
KW   Mouse developmental kinase 1; MDK1 T1; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..28
FT                   /label= sig_peptide
FT   modified_site   64..66
FT                   /label= N-glycosylation_site
FT   modified_site   343..345
FT                   /label= N-glycosylation_site
FT   modified_site   410..412
FT                   /label= N-glycosylation_site
FT   modified_site   555..579
FT                   /label= N-glycosylation_site
FT   domain          555..579
FT                   /label= Transmembrane_domain
FT   misc_difference 600..610
FT                   /note= "product of alternative splicing"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; 128pp; English.
PR   03-JAN-1995; US-368776.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Ciossek T, Millauer B, Ullrich A;
DR   WPI; 96-333988/33.
DR   N-PSDB; T32961.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 109-111; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 truncated versions, MDK1 T1 (W03422) and MDK1 T2
CC   (W03423), of the novel mouse developmental kinase 1 (MDK1) (see also
CC   W03421), a new member of the eck/eph family of receptor tyrosine
CC   kinases. Their amino acid sequences were deduced from cDNA clones
CC   (T32961 and T32962) obtd. from adult mouse brains. MDK T1 and T2
CC   each possess the entire ectodomain, the transmembrane domain and
CC   part of the juxtamembrane region of MDK1, but lack the catalytic
CC   tyrosine kinase domain. They can be used to screen for potential
CC   agents useful for treatment of diseases characterised by abnormal
CC   signal transduction.
SQ   Sequence 610 AA;

Query Match      63.2%;    Score 926.5;    DB 1;    Length 610;
Best Local Similarity   68.8%;   Pred. No. 1.9e-87;
Matches  172;   Conservative 37;   Mismatches 38;   Indels 3;   Gaps 2;

Qy   25  QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT   84
         |||||||  ||| ||| ||||||||||||||:||||||||||||||||||||: |||||||
Db   28  QAAKEVLLLDSKAQQTELEWISSPPSGWEEISGLDENYTPIRTYQVCQVMEPNQNNWLRT   87

Qy   85  NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKID  144
         ||| : |||: ||||||||||||||| ||:|||||||||||: :|:: |||||||||||
Db   88  NWISKGNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLYYETDYDTGRNIRENLYVKID   147

Qy  145  TIAADESFTQMDLGDRILKLNTEIREVGAVYLAFQDVGACVALVSVRVYFKKCPF      204
         ||||||||| :||| |:: ||||:||: |:| |||||||||:|||||:|:|:||:|
Db  148  TIAADESFTQGDLGERKMLNTEVREIGPLSKKGFYLAFQDVGACIALVSVKYYKCWT     207

Qy  205  TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNA  261
         ||:||| ||||| :  ::||||||:|| ::::|   |::::|:||:|||||||||| | |
Db  208  IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEEAENSPRMHCSAEGEWLVPIGKCICKA  267
```

Right column — RESULT 12

```
Qy   262  GYEERGFMCQ  271
Db   268  GYQQKGDTCE  277

RESULT 12
ID   W03423
AC   W03423;
DT   11-NOV-1996  (first entry)
DE   Mouse developmental kinase 1 MDK1 T2.
KW   Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..28
FT                   /label= sig_peptide
FT   modified_site   64..66
FT                   /label= N-glycosylation_site
FT   modified_site   343..345
FT                   /label= N-glycosylation_site
FT   modified_site   410..412
FT                   /label= N-glycosylation_site
FT   modified_site   555..579
FT                   /label= N-glycosylation_site
FT   domain          600..626
FT                   /label= Transmembrane_domain
FT   misc_difference 600..626
FT                   /note= "product of alternative splicing"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; 128pp; English.
PR   03-JAN-1995; US-368776.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Ciossek T, Millauer B, Ullrich A;
DR   WPI; 96-333988/33.
DR   N-PSDB; T32962.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 113-115; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 truncated versions, MDK1 T1 (W03422) and MDK1 T2
CC   (W03423), of the novel mouse developmental kinase 1 (MDK1) (see also
CC   W03421), a new member of the eck/eph family of receptor tyrosine
CC   kinases. Their amino acid sequences were deduced from cDNA clones
CC   (T32961 and T32962) obtd. from adult mouse brains. MDK T1 and T2
CC   each possess the entire ectodomain, the transmembrane domain and
CC   part of the juxtamembrane region of MDK1, but lack the catalytic
CC   tyrosine kinase domain. They can be used to screen for potential
CC   agents useful for treatment of diseases characterised by abnormal
CC   signal transduction.
SQ   Sequence 626 AA;

Query Match      63.2%;    Score 926.5;    DB 1;    Length 626;
Best Local Similarity   68.8%;   Pred. No. 2e-87;
Matches  172;   Conservative 37;   Mismatches 38;   Indels 3;   Gaps 2;

Qy   25  QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT   84
         |||||||  ||| ||| ||||||||||||||:||||||||||||||||||||: |||||||
Db   28  QAAKEVLLLDSKAQQTELEWISSPPSGWEEISGLDENYTPIRTYQVCQVMEPNQNNWLRT   87

Qy   85  NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKID  144
         ||| : |||: ||||||||||||||| ||:|||||||||||: :|:: |||||||||||
Db   88  NWISKGNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLYYETDYDTGRNIRENLYVKID   147

Qy  145  TIAADESFTQMDLGDRILKLNTEIREVGAVYLAFQDVGACVALVSVRVYFKKCPF      204
         ||||||||| :||| |:: ||||:||: |:| |||||||||:|||||:|:|:||:|
Db  148  TIAADESFTQGDLGERKMLNTEVREIGPLSKKGFYLAFQDVGACIALVSVKYYKCWT     207

Qy  205  TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNA  261
         ||:||| |||||  :  ::||||||:|| ::::    |::::|:||:|||||||||| | |
Db  208  IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEEAENSPRMHCSAEGEWLVPIGKCICKA  267
```

Db   208 IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEEAENSPRMICSAEGEWLVPIGKCICKA 267

Qy   262 GYEERGFMCQ 271
              ||::||: ||
Db   268 GYQQKGDTCE 277

RESULT 13

ID   W03421
AC   W03421;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1.
OS   Mus sp.
KW   signal transduction; probe; diagnosis; therapy;
KW   neurodegeneration; neuroproliferation; cancer.
DE   Mouse developmental kinase 1; MDK1; receptor tyrosine kinase; RTK;
FH   Key   Location/Qualifiers
FT   peptide   1..28
FT   /label= Sig_peptide
FT   modified_site 64..66
FT   /label= N-glycosylation_site
FT   modified_site 343..345
FT   /label= N-glycosylation_site
FT   modified_site 410..412
FT   /label= N-glycosylation_site
FT   domain   555..579
FT   /label= Transmembrane_domain
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; 96-333988/33.
PR   03-JAN-1995; US-368776.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Ciossek T, Millauer B, Ullrich A;
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 105-108; 128pp; English.
CC   Mouse developmental kinase 1 (MDK1) (W03421) is a new member of the
CC   eck/eph family of receptor tyrosine kinases (RTKs). Its amino
CC   acid sequence was deduced from a cDNA clone (T32960) isolated
CC   from mouse embryo and adult brain libraries. The distinct
CC   patterns of MDK1 expression during mouse development suggest an
CC   important role for MDK1 in the formation of neuronal structures.
CC   MDK1 may be obtd. by expression in host cells. It can be used
CC   in methods for the diagnosis of diseases characterised by
CC   abnormality in a signal transduction pathway, such as
CC   neuroproliferative or neurodegenerative disorders or cancer,
CC   to screen for (ant)agonists, and to raise antibodies.
SQ   Sequence 998 AA;

Query Match   63.2%; Score 926.5; DB 1; Length 998;
Best Local Similarity   68.8%; Pred. No. 4e-87;
Matches 172; Conservative 37; Mismatches 38; Indels 3; Gaps 2;

Qy   25 QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
              |: ||||| ||||||::|||||||||||||||||||||||||||||||||||||||||||
Db   28 QAAKEVLLLDSKAQQTELEWISSPPSGWEEISGLDENYTPIRTYQVCQVMEPNQNNWLRT 87

Qy   85 NWVPRNSAQKIYVELEKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKID 144
              ||: ::||::|||||||||||||||||||||||||||| |||:| |||||:|| ||:||
Db   88 NWISKGNAQRIFVELEKFTLRDCNSLPGVLGTCKETFNLYIYETDYDTGRNIRENLYVKID 147

Qy   145 TIAADESFTQNDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSRVYFKKCPF 204
              ||||||||||:||||||||||||||||||||||||||||||||||||||:||:|||:|||
Db   148 TIAADESFTQCDLGERKMKLNTEVVREIGPLSKKGFYLAFQDVGACIALVSVRVYKKCWT 207

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNA 261
              |:||||||||| : |||||||||||:||: : |: ||||| ||||||||||||||| |
Db   208 IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEEAENSPRMICSAEGEWLVPIGKCICKA 267

Db   208 IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEEAENSPRMICSAEGEWLVPIGKCICKA 267

Qy   262 GYEERGFMCQ 271
              ||::||: ||
Db   268 GYQQKGDTCE 277

RESULT 14

ID   R75712
AC   R75712;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek5.
KW   Cek5; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis.
OS   Gallus sp.
PN   WO9515375-A.
PD   08-JUN-1995.
PF   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;
DR   WPI; 95-215256/28.
DR   N-PSDB; Q90660.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Disclosure; Page 92-96; 129pp; English.
CC   Probes derived from the Eph-related PTKs Cek4 (Q90669) and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
SQ   Sequence 995 AA;

Query Match   52.7%; Score 772; DB 1; Length 995;
Best Local Similarity   55.4%; Pred. No. 3.8e-71;
Matches 139; Conservative 51; Mismatches 53; Indels 8; Gaps 4;

Qy   29 EVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVP 88
              | :|||:|| ||||||||||::|||||||||| : |||||||||| || |||| ::
Db   28 EETLMDSTTIATAELGWMVHPPSGWEEVSGYDENMNTIRTYQVCNVFESSQNNWLRTKYIR 87

Qy   89 RNSAQKIYVELEKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK----FREHQFTKID 144
              | :||:||||||||||||||:|||||||||||||| |||:||:| :   : |:|||||
Db   88 RGAHRIHVEMKFSVRDCSSIPNVPGSCKETFNLYYESDFDSATKIFPNMWMENPWMKVD 147

Qy   145 TIAADESFTQNDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSRVYFKKCPF 204
              |||||||||||||||||:||:|||| | |||||||||||||||:||||||||:|||
Db   148 TIAADESFSQVDLGGRVNKINIEVRSGPVSKNGFYLAFQDYGGCMSIIAVRVFYRKCPR 207

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAG 262
              :|| || :|||:| :| || ||:||:|:::| |||| : | || |||||||||||| |
Db   208 VIQNGAVFQEILSGAESTSLVAARGTCISNAEEVDVFIKLYCNGDGEWLVPIGRCMCRPG 267

Qy   263 YE--ERGFMCQ 271
              || ||:|:|:
Db   268 YESVENGTVCR 278

RESULT 15

ID   R75709
AC   R75709;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek5+.
KW   Cek5+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis.
OS   Gallus sp.
PN   WO9515375-A.
PD   08-JUN-1995.
PF   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;

```
DR   WPI; 95-215256/28.
DR   N-PSDB; Q90657.
PT   Q90657.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
     cancer.
PS   Claim 11; Page 71-75; 129pp; English.
CC   A cDNA clone encoding a novel variant of EPH-related PTK Cek5,
CC   Cek5+ (Q90657), was isolated from a chick embryo library in
CC   lambda gt11.  Cek5+ protein (R75709) contains a 16-amino acid
CC   insertion in the juxtamembrane domain, and be a result of
CC   alternative splicing.  Cek5+ is exclusively expressed in the CNS.
SQ   Sequence  1011 AA;

Query Match             52.7%;  Score 772;  DB 1;  Length 1011;
Best Local Similarity   55.4%;
Matches  139;  Conservative  51;  Mismatches  53;  Indels   8;  Gaps   4;

Qy   29 EVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVP  88
         |:||||||||||||:|:||||||:|||||||||||||||||:||
Db   28 EETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQVCNVFESSQNNWLRTKYIR  87

Qy   89 RNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK----FREHQFTKID  144
         | :|:||:||||:||::||||   |||||||||||||     :|: ||:|
Db   88 RRGAHRIHVEMKFSVRDCSSIPNVPGSCKETFNLYYYESDFDSATKTFPNWMENPWMKVD  147

Qy  145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVTFKKCPF  204
         ||||||||| || ||||::||:|||||| | |||||||||||||:|||||||||||||
Db  148 TIAADESFSQVDLGGRVMKINTEVRSFGPVSKNGFYLAFQDYGGCMSLIAVRVFYRKCPR  207

Qy  205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAG  262
         ::|| | ||||   |:|| ||||||:||||:|||||  || |||||||||:||||:|
Db  208 VIQNGAVFQEILSGAESTSLVAARGTCISNAEEVDVPIKLYCNGDGEWLVPIGRCMCRPG  267

Qy  263 YE--ERGFMCQ  271
         ||  |:||:|:
Db  268 YESVENGTVCR  278
```

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on:        May 9, 2000, 21:51:35 ; Search time 71.69 Seconds
                                       (without alignments)
                                       212.770 Million cell updates/sec

Title:          US-09-104-340-1
Perfect score:  1200
Sequence:       1 WEEISGVDEHYTPIRIYQVC..........VPIGKCSCNAGYEERGFMCQ 220

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:       225878 seqs, 69334122 residues
Total number of hits satisfying chosen parameters:  225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :  SPTREMBL_12:*
         1: sp_archea:*
         2: sp_bacteria:*
         3: sp_fungi:*
         4: sp_human:*
         5: sp_invertebrate:*
         6: sp_mammal:*
         7: sp_mhc:*
         8: sp_organelle:*
         9: sp_phage:*
        10: sp_plant:*
        11: sp_rodent:*
        12: sp_virus:*
        13: sp_vertebrate:*
        14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	894.5	74.5	985	13	Q91694	xenopus lae
2	894.5	74.5	986	13	Q91845	xenopus lae
3	862.5	71.5	993	13	Q91845	gallus gall
4	838.5	69.9	880	13	073879	brachydanio
5	755.5	63.0	981	13	013146	brachydanio
6	708	59.0	1055	4	043477	homo sapien
7	691.5	57.6	985	13	Q91571	xenopus lae
8	676.5	56.4	943	4	043569	homo sapien
9	676.5	56.4	973	4	095143	homo sapien
10	676.5	56.4	984	4	095142	homo sapien
11	663	56.4	974	13	Q91735	xenopus lae
12	644	53.7	988	13	Q91735	xenopus gall
13	589	49.1	938	11	060669	mus musculu
14	563.5	47.0	902	13	091736	xenopus lae
15	541.5	45.1	952	13	007494	gallus gall
16	509	42.4	976	13	073878	brachydanio
17	464	38.7	1014	11	008644	mus musculu
18	460	38.3	1006	4	015197	homo sapien
19	379.5	31.6	1035	5	Q9XZL6	drosophila
20	379.5	31.6	1080	5	Q9YIJ0	drosophila
21	369.5	30.8	1019	5	096435	drosophila.
22	333.5	27.8	871	13	073875	brachydanio
23	280	23.3	815	13	057458	xenopus lae
24	241	20.1	1122	5	061460	caenorhabdi
25	194.5	16.2	73	13	Q91626	xenopus lae
26	194	16.2	893	5	Q9Y1Y3	ephydatia f
27	89.5	7.5	668	10	065616	arabidopsis
28	88.5	7.4	413	2	005989	staphylococ
29	87.5	7.3	788	13	Q07012	xenopus lae
30	86.5	7.2	305	2	032378	methylobaci
31	86	7.2	919	13	021477	caenorhabdi
32	85.5	7.1	347	5	018395	caenorhabdi
33	84.5	7.0	865	10	023617	arabidopsis
34	84.5	7.0	3110	2	093022	homo sapien
35	82.5	6.9	369	2	050966	borrelia bu
36	81.5	6.8	630	5	013366	caenorhabdi
37	81	6.8	573	2	025431	helicobacte
38	80.5	6.7	684	6	P79101	bos taurus
39	80	6.6	570	2	017698	caenorhabdi
40	79.5	6.6	1436	5	007527	saccharomyc
41	78.5	6.5	1419	5	Q91L5	plasmodium
42	78	6.5	1321	4	Q91115	homo sapien
43	77	6.4	324	5	075129	caenorhabdi
44	77	6.4	806	2	092263	phaedon coc
45	76.5	6.4	336	2	066912	aquifex aeo

ALIGNMENTS

```
RESULT 1
Q91694

ID   Q91694        PRELIMINARY;    PRT;    985 AA.
AC   Q91694;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR PAG PRECURSOR (EC 2.7.1.112)
DE   (PAGLIACCIO).
GN   PAG.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NEURAL CREST;
RX   MEDLINE; 95001564.
RA   WINNING R.S., SARGENT T.D.;
RT   "Pagliaccio, a member of the Eph family of receptor tyrosine kinase
RT   genes, has localized expression in a subset of neural crest and neural
RT   tissues in Xenopus laevis embryos.";
RL   Mech. Dev. 46:219-229(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY
CC       SIMILARITY). MAY PLAY A ROLE IN THE DIFFERENTIATION OF CRANIAL
CC       NEURAL CREST AND OTHER TISSUES.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: LOCALIZED EXPRESSION IN A SUBSET OF NEURAL
CC       CREST AND NEURAL TISSUES IN EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: PRESENT TRANSIENTLY IN VISCERAL ARCH 3. ALSO
CC       EXPRESSED IN THE FOREBRAIN, RHOMBOMERES R3 AND R5 OF THE HINDBRAIN
CC       AND IN THE PRONEPHROS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; L26099; AAA64464.1; -.
DR   HSSP; P00523; 2PRK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V; 1.
```

DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
PFAM; PF00041; fn3; 2.
PFAM; PF00069; pkinase; 1.
PFAM; PF00536; SAM; 1.
PFAM; PF01404; EPH_lbd; 1.
PRINTS; PR00014; FNTYPEIII.
PRINTS; PR00109; TYRKINASE.
PRINTS; PR00104; FNTYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL   1   20   POTENTIAL.
FT   CHAIN   21   985   TYROSINE-PROTEIN KINASE RECEPTOR PAG.
FT   DOMAIN   21   547   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   548   569   POTENTIAL.
FT   DOMAIN   570   985   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   66   115   IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN   191   325   CYS-RICH.
FT   DOMAIN   326   436   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   436   533   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   620   881   PROTEIN KINASE.
FT   DOMAIN   626   634   ATP (BY SIMILARITY).
FT   NP_BIND   626   634   ATP (BY SIMILARITY).
FT   DISULFID   73   108   BY SIMILARITY.
FT   BINDING   652   652   ATP (BY SIMILARITY).
FT   ACT_SITE   745   745   BY SIMILARITY.
FT   BINDING   652   652   BY SIMILARITY.
FT   MOD_RES   778   778   PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD   340   340   POTENTIAL.
FT   CARBOHYD   407   407   POTENTIAL.
SQ   SEQUENCE   985 AA; 109729 MW; 026BA8A5 CRC32;

Best Local Similarity   74.5%;   Score 894.5; DB 13; Length 985;
Query Match   71.9%;   Pred. No 1.8e-79;
Matches 159; Conservative 33; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||:| |||||||||||||||||||||||||||||:|||||:|||||||||||||||||
Db   54 WEEVSIMDEKNTPIRTYQVCNVMESSQNNWLRTDWIPRSGAQRVYVEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          :|||||||||||||||:|| ||:||||||||:|||||||||||: ||||:|||||||:||
Db   114 VMGTCKETFNLYYESSNNDKERFIRETQYVKIDTIAADESFTQVDIGDRIMKLNTEVRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVFVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          |||:|||||||||||||||:|||||||||||||||||:|||||:||| |:|||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYFKKCPLTVRNLAQFPDTITGSDTSSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||| :| :||| || |||||:|||: | ||||||||||
Db   234 VDNSEEKDVPKMYCGADGEWLVPIGNCLCNAGFEEHNGGCQ 274

RESULT   2

Q91845

ID   Q91845   PRELIMINARY;   PRT;   986 AA.
AC   Q91845;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR SEK-1 PRECURSOR (EC 2.7.1.112).
DE   SEK-1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96125143.
RA   XU Q., ALLDUS G., HOLDER N., WILKINSON D.G.;
RT   "Expression of truncated Sek-1 receptor tyrosine kinase disrupts the
RT   segmental restriction of gene expression in the Xenopus and zebrafish
RT   hindbrain.";
RL   Development 121:4005-4016(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY

CC   SIMILARITY). REQUIRED FOR INTERACTIONS THAT REGULATE THE SEGMENTAL
CC   IDENTITY OR MOVEMENT OF CELLS.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSION OCCURS IN R3, R5 AND TRANSIENTLY
CC   AT LOWER LEVELS IN R2.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; X91191; CAA62601.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL   1   15   POTENTIAL.
FT   CHAIN   16   986   TYROSINE-PROTEIN KINASE RECEPTOR SEK-1.
FT   DOMAIN   16   547   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   548   569   POTENTIAL.
FT   DOMAIN   570   986   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   66   115   IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN   191   325   CYS-RICH.
FT   DOMAIN   326   435   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   436   533   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   620   881   PROTEIN KINASE.
FT   NP_BIND   626   634   ATP (BY SIMILARITY).
FT   DISULFID   73   108   BY SIMILARITY.
FT   BINDING   652   652   ATP (BY SIMILARITY).
FT   ACT_SITE   745   745   BY SIMILARITY.
FT   MOD_RES   778   778   PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD   340   340   POTENTIAL.
FT   CARBOHYD   407   407   POTENTIAL.
SQ   SEQUENCE   986 AA; 109840 MW; 028DBBDA CRC32;

Query Match   71.9%;   Score 894.5; DB 13; Length 986;
Best Local Similarity   74.5%;   Pred. No 1.9e-79;
Matches 159; Conservative 33; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||:| |||||||||||||||||||||||||||||:|||||:|||||||||||||||||
Db   54 WEEVSIMDEKNTPIRTYQVCNVMESSQNNWLRTDWIPRSGAQRVYVEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          :|||||||||||||||:|| ||:||||||||:|||||||||||: ||||:|||||||:||
Db   114 VMGTCKETFNLYYESSNNDKERFIRETQYVKIDTIAADESFTQVDIGDRIMKLNTEVRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVFVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          |||:|||||||||||||||:|||||||||||||||||:|||||:||| |:|||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYFKKCPLTVRNLAQFPDTITGSDTSSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||| :| :||| || |||||:|||: | ||||||||||
Db   234 VDNSEEKDVPKMYCGADGEWLVPIGNCLCNAGFEEHNGGCQ 274

RESULT   3

042422

ID   042422   PRELIMINARY;   PRT;   993 AA.
AC   042422;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE-
DE   PROTEIN KINASE RECEPTOR CEPHA7).
GN   CEPHA7.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 98092111.
RA   ARAUJO M., NIETO M.A.;
RT   "The expression of chick Epha7 during segmentation of the central and
RT   peripheral nervous system.";
RL   Mech. Dev. 68:173-177(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WITHIN THE NERVOUS SYSTEM, EXPRESSION IS
CC       RESTRICTED TO PROSOMERES 1 AND 2 IN THE DIENCEPHALON AND ALL THE
CC       RHOMBOMERES IN THE HINDBRAIN DURING SEGMENTATION STAGES. LATER ON,
CC       A SUPERIMPOSED PATTERN APPEARS THAT CORRELATES WITH THE FORMATION
CC       OF SEVERAL AXONAL TRACTS. IN THE SOMITIC MESODERM, THE EXPRESSION
CC       CORRELATES WITH SEGMENTATION AND THE GUIDANCE OF BOTH NEURAL CREST
CC       AND MOTOR AXONS THROUGH THE SCLEROTOMES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; Y14271; CAA74643.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    993       EPH-LIKE RECEPTOR TYROSINE KINASE.
FT   DOMAIN       31    551       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    552    572       POTENTIAL.
FT   DOMAIN      573    993       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       67    116       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      192    328       CYS-RICH.
FT   DOMAIN      329    438       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      439    536       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      628    889       PROTEIN KINASE.
FT   NP_BIND     634    642       ATP (BY SIMILARITY).
FT   DISULFID     74    109       BY SIMILARITY.
FT   BINDING     660    660       ATP (BY SIMILARITY).
FT   ACT_SITE    753    753       BY SIMILARITY.
FT   MOD_RES     786    786       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    343    343       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
SQ   SEQUENCE   993 AA;  111366 MW;  6B6BFD17 CRC32;
```

```
Query Match             71.9%;  Score 862.5;  DB 13;  Length 993;
Best Local Similarity   69.5%;  Pred. No. 2.6e-76;
Matches  155;  Conservative  39;  Mismatches  26;  Indels  3;  Gaps  2;

Qy    61 VLGTCKETFNLIYMESDDHGVKFREHOFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||| :||| :|| ||| :||||||||||||||| |||||||:||||:||
Db   115 VLGTCKETFNLIYYETDYDTGRNIRENQVKIDTIAADESFTQGDLGERMKLNTEVREI 174

Qy    55 WEEISGLDENYTPIRTYQVCNVWESNQNNWLRTHWIARKSNAQRIFVELKFTLRDCNSLPG 114
         |||||| ||||||||||||||| :||:||||||:|| :||:||| |:|||||||||||:| :|
Db   121 GPVNKKGFYILAFQDVGACVALVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179

Qy     1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||||| |: ||||||||||||| |:||||||| |: :|:||||:|||||||||||||:|
Db   175 GPLSKKGFYILAFQDVGACIALVSVKVYIKKCWSIIENLAIFPDIVTGSEFSSLVEVRGTC 234

Db   180 VNNSKE--EDPPRMYCSTEGEWLVPIGKGCSCNAGYEERGFMCQ 220
         :: |||    ||| ||||| ||||||||| | ||||:|||| :||
Db   235 VSSAEEEAENSPKMHCSAEGEWLVPIGKGICKCAGYQQKGDTCE 277
```

```
RESULT 4
073879
ID   073879;                                  PRELIMINARY;          PRT;   880 AA.
AC   073879;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE RTK4 (FRAGMENT).
GN   RTK4.
OS   Brachydanio rerio (zebrafish) (zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC   Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC   Cyprinoidea; Cyprinidae; Rasborinae; Danio.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   COOKE J.E.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   COOKE J.E., XU Q., WILSON S.W., HOLDER N.;
RL   Dev. Genes Evol. 206:515-531(1997).
RL   EMBL; AJ005030; CAA06303.1; -.
ZFIN; ZDB-GENE-990415-61; rtk4.
HSSP; P00523; 2PTK.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00041; fn3; 2.
FT   NON_TER       1      1
FT   NON_TER       1      1
SQ   SEQUENCE   880 AA;  98862 MW;  7C139606 CRC32;
```

```
Query Match             69.9%;  Score 838.5;  DB 13;  Length 880;
Best Local Similarity   68.0%;  Pred. No. 5e-74;
Matches  149;  Conservative  36;  Mismatches  33;  Indels  1;  Gaps  1;

Qy     3 EISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVL 62
         :|| :|| ||||||||||||| :|:|||||||:|| :|:|||||:|:||||||||||:| :
Db     1 EVSIMDEKNIPIRTYQVCNVMEPSQNNWLRTHWIORGAAQRIVIEIKFTLRDCNSLPGVI 60

Qy    63 GTCKETFNLIYLESDDHGVKFREHOFTKIDTIAADESFTQMDLGDRILKLNTEIREVGP 122
         |||||||||||||:|| :| :||:|| :||||||||||||:|||||||:|||||:||:|
Db    61 GTCKETFNLIYLESDDNERYAHESRFSKIDTVAADESFTQVDLGDRIMKLNTEVRDVGV 120

Qy   123 VNRKGFYILAFQDVGACVALVSRVYTFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVN 181
         :|| ||||||||||||:|||||: ||:|||| ||:|||| |||||| ||:::||||||||:
Db   121 LSRAGFYILAFQDVGACIALVSVHVFYIKKCPLAVRNLAQPPDTVTGADTSSLVEVRGSCVD 180

Qy   182 NSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         :|: : ||||||| ||||||||| | || ||||||| :||
Db   181 HSEEQEVPRKMYCCADGEWLVPIGNCLCNPGYEERNEQCQ 219
```

```
RESULT 5
013146
ID   013146;                                  PRELIMINARY;          PRT;   981 AA.
AC   013146;
```

DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE KINASE 1 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN KINASE
DE   ZEK1).
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC   Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC   Cyprinoidea; Cyprinidae; Rasborinae; Danio.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97329578.
RA   BOVENKAMP D.E., GREER P.;
RT   "Novel Eph-family receptor tyrosine kinase is widely expressed in the
RT   developing zebrafish nervous system.";
RL   Dev. Dyn. 209:166-181(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC   ROLE IN EARLY PATTERN FORMATION WITHIN THE DEVELOPING NERVOUS
CC   SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + A PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN THE DEVELOPING ZEBRAFISH
CC   NERVOUS SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPH RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U89295; AAC60220.1; -.
DR   HSSP; P05523; 2PTK.
DR   ZFIN; ZDB-GENE-990415-58; zek1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; FALSE_NEG.
DR   PFAM; PF000069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   20   BY SIMILARITY.
FT   CHAIN   21   981   EPH-LIKE KINASE 1.
FT   DOMAIN   21   545   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   546   566   POTENTIAL.
FT   DOMAIN   567   981   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   64   113   IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN   192   325   CYS-RICH.
FT   DOMAIN   326   438   FIBRONECTIN TYPE-III.
FT   DOMAIN   439   532   FIBRONECTIN TYPE-III.
FT   DOMAIN   626   887   PROTEIN KINASE.
FT   NP_BIND   632   640   ATP (BY SIMILARITY).
FT   DISULFID   71   106   BY SIMILARITY.
FT   BINDING   658   658   ATP (BY SIMILARITY).
FT   ACT_SITE   751   751   BY SIMILARITY.
FT   MOD_RES   784   784   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD   340   340   POTENTIAL.
FT   CARBOHYD   410   410   POTENTIAL.
FT   CARBOHYD   435   435   POTENTIAL.
FT   CARBOHYD   485   485   POTENTIAL.
FT   VARIANT   141   141   S -> N.
SQ   SEQUENCE   981 AA; 109654 MW; 10D38182 CRC32;

Query Match   63.0%; Score 755.5; DB 13; Length 981;
Best Local Similarity   61.4%;
Matches 137; Conservative 37; Mismatches 46; Indels   3; Gaps   1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
              ||||| ||||:||| ||||||||:||::|||| |||:|:::||:||||||||||||||||:|
Db   52 WEEISVMDERNIPMRTYQVCNVMEANQNNWLRTGLIQREGAQRYVEIKFTLRDCNSLPG 111

Qy   61 VLGTCKETFNLYYMESDDDHGV--KFREHQFTKIDTIAADESFTQMDLGDRILKLNTEI 117
              |||||||||||| :||| :: | ||:||||||||||||||||||||||: | |||||:|
Db   112 VPGTCKETFNVYIHESNNAVAAPLRHIRESQYIKIDTIAADESFTQTDVGDRVNKLNTEV 171

Qy   118 REVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRG 177
              |::| |||||||||||: ||||||||:|||||||||||||||||||||:|||::|||||||
Db   172 RDISGLSKRGLYLAFQDLGACIALVSVRVFYKRCPLAVLNLARPPDTVTGGDSALVEVRG 231

Qy   178 SCVNNSKEDDPRMYCSTEGEWLYPIGKCSCNAGYEERGFMCQ 220
              ||::||:::||| :|||:|||::|| |:|:|||||::|||:||
Db   232 TCVEDAEELEGPRMFCSADGGWLVPIGRCVCRPGFEEVDGHCQ 274

RESULT   6
ID   O43477   PRELIMINARY;   PRT;   1055 AA.
AC   O43477;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTEIN-TYROSINE KINASE EPHB2V.
GN   EPHB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96154673.
RA   IKEGAKI N., TANG X.X., LIU X.G., BIEGEL J.A., ALLEN C., YOSHIOKA A.,
RA   SULMAN E.P., BRODEUR G.M., PLEASURE D.E.;
RT   "Molecular characterization and chromosomal localization of DRT
RT   (EPHT3): a developmentally regulated human protein-tyrosine kinase
RT   gene of the EPH family.";
RL   Hum. Mol. Genet. 4:2033-2045(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   TANG X.X., PLEASURE D.E., IKEGAKI N.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF025304; AAB94602.1; -.
DR   HSSP; P05523; 2PTK.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF000069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Tyrosine-protein kinase.
SQ   SEQUENCE 1055 AA; 117492 MW; EAF28155 CRC32;

Query Match   59.0%; Score 708; DB 4; Length 1055;
Best Local Similarity   55.7%;
Matches 127; Conservative 49; Mismatches 44; Indels   8; Gaps   4;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
              |||||| ||:||| ||||||||||:||:: | ||::::::||:||||||||||||||||||
Db   43 WEEVSGYDENMNTIRTYQVCNVFESSQNNWLRTKFIRRGAHRIHVEMKFSVRDCSSIPS 102

Qy   61 VLGTCKETFNLYYMESDDDHGVK------FREHQFTKIDTIAADESFTQMDLGDRILKLNTE 116
              ||||||||||| |||:::| ||: ||::||||||||||||:||||||||||||:||
Db   103 VPGSCKETFNLYYIYEADFDSAIKTFPNWMENPWVKVDTIAADESFSQVDLGGRVNKINTE 162

Qy   117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
              ||||||||||||||||||:|||::|||:||||||||||:||:||||||||| ||:|||:|
Db   163 VRSFGPVSRSGFYLAFQDYGGCMSLIAVRVFYRKCPRIIONGAIFQETLSGAESTSLVAA 222

Qy   176 RGSCVNNSKEDPP--RMYCSTEGEWLVPIGKCSCNAGYE--ERGFMCQ 220

RESULT 7

Db   223 RGSCIANAEEVDVPIKLYCNGDGEWLVPIGRCMCKRAGFEAVENGTVCR 270
          |||:|| | |||:| | | ||||||||||||| | ||| || ||:|:
          |||:|| | |||:| |:|| : ||||| |||||||| || || |::

ID   Q91571;   PRELIMINARY;   PRT;   985 AA.
AC   Q91571;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR XEK PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR XEK).
GN   XEK.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95215070.
RA   JONES T.L., KARAVANOVA I., MAENO M., ONG R.C., KUNG H.F., DAAR I.O.;
RT   "Expression of an amphibian homolog of the Eph family of receptor
RT   tyrosine kinases is developmentally regulated.";
RL   Oncogene 10:1111-1117(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC   ROLE IN THE DEVELOPMENT OR FUNCTION OF THE CENTRAL NERVOUS SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED, IT DECREASES AT MID
CC   BLASTULA TRANSITION AND REAPPEARS AT LATE NEURULATION. EXPRESSED
CC   AT HIGHER LEVELS IN THE ANTERIOR AND DORSAL REGIONS OF EMBRYONIC
CC   STAGES 16, 24 AND 37. IN ADULT IT APPEARS TO BE UBIQUITOUSLY
CC   EXPRESSED WITH HIGHER EXPRESSION IN BRAIN AND OVARY. EXPRESSION IN
CC   THE BRAIN, BRACHIAL ARCHES, TRIGEMINAL FACIAL GANGLION, AND THE
CC   RETINA OF SWIMMING TADPOLE STAGE OF DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U14164; AAA74888.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; FALSE_NEG.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00109; TYRKINASE.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   19   POTENTIAL.
FT   CHAIN   20   985   EPHRIN TYPE-B RECEPTOR XEK.
FT   DOMAIN   6   12   POLY-LEU.
FT   DOMAIN   20   542   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   543   563   POTENTIAL.
FT   DOMAIN   564   985   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   56   105   IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN   185   321   CYS-RICH.
FT   DOMAIN   322   431   FIBRONECTIN TYPE-III.
FT   DOMAIN   432   529   FIBRONECTIN TYPE-III.
FT   DOMAIN   620   883   PROTEIN KINASE.
FT   NP_BIND   626   634   ATP (BY SIMILARITY).
FT   DISULFID   63   98   BY SIMILARITY.
FT   BINDING   652   652   ATP (BY SIMILARITY).
FT   ACT_SITE   745   745   BY SIMILARITY.
FT   MOD_RES   779   779   PHOSPHORYLATION (BY SIMILARITY).
FT   CARBOHYD   336   336   POTENTIAL.
FT   CARBOHYD   428   428   POTENTIAL.
FT   CARBOHYD   482   482   POTENTIAL.
SQ   SEQUENCE   985 AA; 110104 MW; 24166A8F9 CRC32;

Query Match   57.6%; Score 691.5; DB 13; Length 985;
Best Local Similarity   55.1%; Pred. No. 1.5e-59;
Matches   125; Conservative   47; Mismatches   48; Indels   7; Gaps   4;

Qy   1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
          ||||:|| ||||:|| |:||| || : ||||: ||| :||||:| ||||||||| |||
Db   44 WEEVSGVDENLNTIRIYQVCNVFGPKQNNWLLTTFIPRRGAIRVYVEMRFTVRDCSSLPN 103

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTE 116
          ||||||||||||||||:| ||| :|||| ||||||||| |||| :|:||||||:|||
Db   104 VPGSCKETFNLYYETDSNIENKISIFWNESPYLKVDTIIAADESFSQVDFGGRLMKVNTE 163

Qy   117 IREVGPVNKKGFYLAFQDVGACVALVSRVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
          ||| || |||||:||||| |||::|:|||||||||||:: |:|:| |: : |||||||
Db   164 VRSFGPLTRSGFYLAFQDYGACMSLLSVRVFFKEMPSSVVQNLLVFPETMTGAESTSLVIA 223

Qy   176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ 220
          |:|: |||:|| | :||||||| ||||| ||| |||| ||| ||
Db   224 RGTCIPNAEEVDVPIKLYCNGDGEWMVPIGKCTCKAGYEPENHVVCK 270

RESULT 8

ID   O43569;   PRELIMINARY;   PRT;   943 AA.
AC   O43569;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1.
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1C
DE   (EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1D).
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   STEIN E., SCHOECKLMANN H.O., DANIEL T.O.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF037333; AAB94627.1; -.
DR   EMBL; AF037334; AAB94628.1; -.
DR   HSSP; P08631; 2HCK.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 2.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00014; FNTYPEIII.
SQ   SEQUENCE   943 AA; 105291 MW; C4C78A21 CRC32;

Query Match   56.48; Score 676.5; DB 4; Length 943;
Best Local Similarity   54.28; Pred. No. 4.3e-58;
Matches   123; Conservative   45; Mismatches   52; Indels   7; Gaps   4;

Qy   1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||:||:|||:|| |:|||||::|| |||| |||| ||| |:||||||||||||||||:
Db   42 WEEVSGYDENLNTIRIYQVCNVFEPNQNNWLLTTFINRRGAIRIYSEMRFTVRDCSSLPN 101

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTE 116
          ||||||||||||||||:| ||| :|||| ||||||||| |||| :|:||||||:|||
Db   102 VPGSCKETFNLYYETDSVIATKKSAPWSEAPYLKVDTIAADESFSQVDFRGRLMKVNTE 161

Qy   117 IREVGPVNKKGFYLAFQDVGACVALVSRVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
          ||| || |||||:|||||:|||::|:|||||::|||: :|:|||| |: :| |||| ||
Db   162 VRSFGPLTRNGFYLAFQDYGACMILLSVRVFFKKCPSIVQNFAVFPETMTGAESTSLVIA 221

Qy   176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ 220
          |:|: |||:|| | :||||||| |||||| ||| |||| ||| ||
Db   162 VRSFGPLTRNGFYLAFQDYGACMILLSVRVFFKKCPSIVQNFAVFPETMTGAESTSLVIA 221

Db   222 RGTCIPNAEEVDVPIKLYCNGDGEWWPIGRCTCKPGYEPENSVACK 268

RESULT   9
ID   O95143   PRELIMINARY;   PRT;   973 AA.
AC   O95143;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1B.
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY.
RA   SCHOECKLMANN H.O., DANIEL T.O.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF037332; AAD02031.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Receptor; Kinase.
SQ   SEQUENCE   973 AA; 108938 MW; 548C281A CRC32;

Query Match   56.4%; Score 676.5; DB 4; Length 973;
Best Local Similarity   54.2%; Pred. No. 4.5e-58;
Matches 123; Conservative   45; Mismatches   52; Indels   7; Gaps   4;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||:||:||| |:||||||||||| |||:|||||:|| ||||||||:|:|||||||||
Db   31 WEEVSGYDENLNTIRTYQVCNVFEPNQNNWLLTFINRGAHRIYSEMRFTVRDCSSLPN 90

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDIIAADESFTQMDLGDRILKLNTE 116
         | :||||||||| :| || :|||:|||||||||||||:|||:|| ||:|:|| |
Db   91 VPGSCKETFNLYYETDSVIATKKSAFWSEAPYLKVDTIAADESFSQVDFRGRLMKVNTE 150

Qy 117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         :| :||:|:|||||||||:|||::|:|||||:|||| :|:| |:|| :| :|| ||| |
Db 151 VRSFGPLTRNGFYLAFQDYGACMILLSVRVFKKCPSIVQNFAVFPETMTGAESTSLVIA 210

Qy 176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ 220
         ||:|::|:|| | | ||||:| |||::|||:|:|:||| |:||||:| |
Db 211 RGTCIPNAEEVDVPIKLYCNGDGEWWVPIGRCTCKPGYEPENSVACK 257

RESULT   10
ID   O95142   PRELIMINARY;   PRT;   984 AA.
AC   O95142;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY.
RA   STEIN E., HUYNH-DO U., LANE A., CERRETTI D.P., DANIEL T.O.;
RT   "Nck recruitment to Eph receptor, EphB1/ELK, couples ligand activation
RT   to c-Jun kinase.";
RL   J. Biol. Chem. 0:0-0(1997).
EMBL; AF037331; AAD02030.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Receptor; Kinase.

Query Match   56.4%; Score 676.5; DB 4; Length 984;
Best Local Similarity   54.2%; Pred. No. 4.6e-58;
Matches 123; Conservative   45; Mismatches   52; Indels   7; Gaps   4;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||:||:||| |:||||||||||| |||:|||||:|| ||||||||:|:|||||||||
Db   42 WEEVSGYDENLNTIRTYQVCNVFEPNQNNWLLTFINRGAHRIYSEMRFTVRDCSSLPN 101

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDIIAADESFTQMDLGDRILKLNTE 116
         | :||||||||| :| || :|||:|||||||||||||:|||:|| ||:|:|| |
Db 102 VPGSCKETFNLYYETDSVIATKKSAFWSEAPYLKVDTIAADESFSQVDFRGRLMKVNTE 161

Qy 117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         :| :||:|:|||||||||:|||::|:|||||:|||| :|:| |:|| :| :|| ||| |
Db 162 VRSFGPLTRNGFYLAFQDYGACMILLSVRVFKKCPSIVQNFAVFPETMTGAESTSLVIA 221

Qy 176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ 220
         ||:|::|:|| | | ||||:| |||::|||:|:|:||| |:||||:| |
Db 222 RGTCIPNAEEVDVPIKLYCNGDGEWWVPIGRCTCKPGYEPENSVACK 268

SQ   SEQUENCE   984 AA; 110041 MW; DF5C8ED3 CRC32;

RESULT   11
ID   Q91735   PRELIMINARY;   PRT;   974 AA.
AC   Q91735;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR) (TCK).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96088901.
RA   SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;
RT   "Novel members of the eph receptor tyrosine kinase subfamily expressed
RT   during Xenopus development.";
RL   Oncogene 11:1745-1752(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY (BY
CC   SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO IN PRE-SOMITIC
CC   MESODERM, CAUDAL SOMITES, MIDBRAIN, AND CEMENT GLAND. MOST
CC   ABUNDANT IN ADULT BRAIN, EYE, HEART, LUNG AND OVARY. LOWER LEVELS
CC   IN INTESTINE, KIDNEY, OVIDUCT AND PHARYNX.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L43620; AAA35526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   16   POTENTIAL.

Query Match   56.4%; Score 676.5; DB 4; Length 984;
Best Local Similarity   54.2%; Pred. No. 4.6e-58;
Matches 123; Conservative   45; Mismatches   52; Indels   7; Gaps   4;

```
FT   CHAIN           17    974       EPHRIN TYPE-B RECEPTOR.
FT   DOMAIN          17    534       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       535    555       POTENTIAL.
FT   DOMAIN         556    974       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN          53    102       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN         178    315       CYS-RICH.
FT   DOMAIN         316    423       FIBRONECTIN TYPE-III.
FT   DOMAIN         424    520       FIBRONECTIN TYPE-III.
FT   DOMAIN         609    872       PROTEIN KINASE.
FT   DOMAIN         615    623       ATP (BY SIMILARITY).
FT   DISULFID        60     95       BY SIMILARITY.
FT   BINDING        641    641       ATP (BY SIMILARITY).
FT   BINDING        734    734       ATP (BY SIMILARITY).
FT   NP_BIND        615    623       BY SIMILARITY.
FT   ACT_SITE       768    768       BY SIMILARITY.
FT   MOD_RES        768    768       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD       330    330       POTENTIAL.
FT   CARBOHYD       420    420       POTENTIAL.
SQ   SEQUENCE       974 AA;  108263 MW;  BD419F6D CRC32;
```

```
Query Match                        55.2%;  Score 663;  DB 13;  Length 974;
Best Local Similarity   53.5%;  Pred. No. 9.5e-57;
Matches  122;  Conservative  41;  Mismatches  53;  Indels  12;  Gaps   5;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         ||||:||| ||||||||||||||| :|||||||||||||||| :|||||||||||||:|:|
Db   41 WEEVSGYDEASNPIRTYQVCNVRDSNQNNWLRTQFIPRQDVQRVVELKFTVRDCNSLPN 100

Qy   61 VLGTCKETFNLIYYMESDDDHGVK----FREHQFTKIDIIAADESFTQMDLGDRILKLNTE 116
         :|:|||||||| |:| ||||  ||  | : ||||||||||||||:|||||||||||||||:|
Db  101 LRGSCKETFNFFYYESDSDSASADSPFWMENPYIKVDTIAPDESFSRRDSG----RVNTK 156

Qy  117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         |:||||||||||||||||||:||||||||||||||||||||||||||||| ::|||:
Db  157 IRSFGPISRAGFYLAFQDLGDACVSLISVRVFFKKCPRTRAGFAASPPETITGAEPTSLVIA 216

Qy  176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYEERG--FMCQ 220
         ||||||||||||||| :||::|||||||::||:|||||||| ||:||
Db  217 PGTCVPNALEVSVPLKIYCNGDGDWMVPVGACTCAAGFEPAGKDTQCQ 264
```

```
RESULT  12
ID      Q07498;      PRELIMINARY;      PRT;      988 AA.
AC      Q07498;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE      TYROSINE KINASE CEK10 RECEPTOR (EC 2.7.1.112) (FRAGMENT).
GN      CEK10.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC      Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=EMBRYO;
RX      MEDLINE; 93288394.
RA      SAJADI F.G., PASQUALE E.B.;
RT      "Five novel avian Eph-related tyrosine kinases are differentially
RT      expressed.";
RL      Oncogene 8:1807-1813(1993).
DR      EMBL; Z19061; CAA79511.1; -.
DR      HSSP; P00523; 2PTK.
DR      PROSITE; PS00523; 2PTK.
DR      PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR      PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PFAM; PF01404; EPH_lbd; 1.
DR      PFAM; PF00041; fn3; 2.
DR      PFAM; PF00536; SAM; 1.
DR      PFAM; PF00069; pkinase; 1.
KW      Transferase.
SQ      SEQUENCE       988 AA;  109578 MW;  C5CA9B7B CRC32;
```

```
Query Match                        53.7%;  Score 644;  DB 13;  Length 988;
Best Local Similarity   51.3%;  Pred. No. 7.1e-55;
Matches  117;  Conservative  48;  Mismatches  51;  Indels  12;  Gaps   5;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         ||||:||| ||||||||||||||| ::||||||||| :|||||||||||||||:||:||:|
Db   34 WEEVSGYDEAMNPIRTYQVCNVREANQNNWLRTKFIQRQDVQRVVELKFTVRDCNSIPN 93

Qy   61 VLGTCKETFNLIYYMESDDD----HGVKFREHQFTKIDIIAADESFTQMDLGDRILKLNTE 116
         :|:|||||||| |||||||| |   :|: :|||||||||||||||:||||||||||||:
Db   94 IPGSCKETFNLFYYESDTDSASANSPFWMENPYIKVDTIAPDESFSKLESG----RVNTK 149

Qy  117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         :|||||||||||||||||||:||||||||||||||:|||||||||||||| ::|||:
Db  150 VRSFGPLSKNGFYLAFQDLGDACMSLISVRAFYKKCSNTIAGFAIFPETLTGAEPTSLVIA 209

Qy  176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYEE--RGFMCQ 220
         |||||||||||||| :|||::|||||||::||:|||||| ||:||
Db  210 PGTCIPNAVEVSVPLKIYCNGDGEWMVPVGACTCAAGYEPAMKDTQCQ 257
```

```
RESULT  13
ID      Q60669;      PRELIMINARY;      PRT;      938 AA.
AC      Q60669;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE      PROTEIN KINASE (FRAGMENT).
GN      EPHB3 OR ETK2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE; 95200798.
RA      RUIZ J.C., CONLON F.L., ROBERTSON E.J.;
RT      "Identification of novel protein kinases expressed in the myocardium
RT      of the developing mouse heart.";
RL      Mech. Dev. 48:153-164(1994).
DR      EMBL; U11493; AAA67925.1; -.
DR      HSSP; P00523; 2PTK.
DR      MGD; MGI:104770; Ephb3.
DR      PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR      PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PFAM; PF01404; EPH_lbd; 1.
DR      PFAM; PF00041; fn3; 2.
DR      PFAM; PF00536; SAM; 1.
DR      PFAM; PF00069; pkinase; 1.
NON_TER        1      1
SQ      SEQUENCE       938 AA;  103996 MW;  2B01D2A0 CRC32;
```

```
Query Match                        49.1%;  Score 589;  DB 11;  Length 938;
Best Local Similarity   49.8%;  Pred. No. 1.7e-49;
Matches  108;  Conservative  45;  Mismatches  54;  Indels  10;  Gaps   4;

Qy    3 EISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVL 62
         |:|| || |||||||||||| |||||||||| :: :|||||||||||||||:||:||:|:
Db    1 EVSGYDEAMNPIRTYQVCNVRESSQNNWLRTGFIWRREVQRVYVELKFTVRDCNSIPNIP 60

Qy   63 GTCKETFNLIYYMESDDD----HGVKFREHQFTKIDIIAADESFTQMDLGDRILKLNTEIR 118
         ||||||||| :|||||| |   :|: :|||||||||||||||:||||||||||||:||
Db   61 GSCKETFNLFYYEADSDVASASSPFWMENPYVKVDDISPEDSFSRSMPG----RVNTKVR 116

Qy  119 EVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRG 177
         ||||||||||||||||||:|||||||||:|||||||:||||||||||| ::|||:
Db  117 SFGPLSKAGFYLGFQOGACMSLISVRAFYKKCASTIAGFALFPELTGAEATSLVIAPR 176
```

```
Qy  178 SCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE 213
            :|: ||| |:|:||||:|:|||:||:||| |:|
Db  177 ACIANAVEVSVPLKLYCNGDGEWMVPVGACTCATGHE 213
```

```
RESULT 14

ID   Q91736          PRELIMINARY;   PRT;   902 AA.
AC   Q91736;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH RECEPTOR TYROSINE KINASE (XELK) (FRAGMENT).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96068901.
RA   SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;
RT   "Novel members of the eph receptor tyrosine kinase subfamily expressed
RT   during Xenopus development.";
RL   Oncogene 11:1745-1752(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO IN THE BRAIN AND
CC       SPINAL CORD AND IN THE FIRST AND FOURTH VISCERAL ARCHES. MOST
CC       ABUNDANT IN ADULT BRAIN, WITH LOWER LEVELS IN EYE, HEART, OVARY,
CC       OVIDUCT, LUNG AND PHARYNX.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L43621; AAA93527.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmembrane; Glycoprotein; Immunoglobulin domain;
KW   Transferase; Tyrosine-protein kinase; Atp-binding; Phosphorylation;
FT   NON_TER        1      1
FT   DOMAIN        <1    459
FT   TRANSMEM     460    480       POTENTIAL.
FT   DOMAIN       481    902       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        <1     21       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       101    237       IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN       238    347       CYS-RICH.
FT   DOMAIN       348    445       FIBRONECTIN TYPE-III.
FT   DOMAIN       537    800       PROTEIN KINASE.
FT   DOMAIN       543    551       FIBRONECTIN TYPE-III.
FT   NP_BIND      543    551       ATP (BY SIMILARITY).
FT   DISULFID       ?     14       BY SIMILARITY.
FT   BINDING      569    569       ATP (BY SIMILARITY).
FT   ACT_SITE     662    662       BY SIMILARITY.
FT   MOD_RES      696    696       PHOSPHORYLATION.
FT   CARBOHYD     252    252       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD     252    252       POTENTIAL.
FT   CARBOHYD     344    344       POTENTIAL.
FT   CARBOHYD     398    398       POTENTIAL.
SQ   SEQUENCE    902 AA;   100850 MW;   C09D6D06 CRC32;
```

```
Query Match              47.0%;   Score 563.5;   DB 13;   Length 902;
Best Local Similarity    55.1%;   Pred. No. 5e-47;
Matches 102;  Conservative 42;  Mismatches 34;  Indels 7;  Gaps 4;
```

```
Qy   43 KIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK-----FREHQFTKIDIIAAD  98
         ::||||:|||:|||:|| | :|||||||||| :| ||      :|    |||||:|||||||
Db    2 RVYVEMRFTVRDCSSLPNVPGSCKETFNLYYETDSNIDNKISTFWNESPYLKVDIIAAD   61

Qy   99 ESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRYFKKCPFTVKNL  158
         |||:| | | |::|:|||:|   ||: ||||||||| |||::|:|||  ||||| :|:|
Db   62 ESFSQVDFGGRLMKVNTEVRSPGPLTRSGFYLAFQDYGACMSLLSVRFVFFKKCPSVVQNF 121

Qy  159 AMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ER  215
         | || :   :|| ||:  || |:   :|| ||    |: ||:||||||||:| ||| |
Db  122 AVFPEIMTGAESTSLVIARGTCIPNAEEVDVPIKLYCNGDGEWMVPIGKCTCKAGYEPEN 181

Qy  216 GFMCQ  220
          ::|:
Db  182 HVVCK  186
```

```
RESULT 15

ID   Q07494          PRELIMINARY;   PRT;   952 AA.
AC   Q07494;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CEK6 PROTEIN (FRAGMENT).
GN   CEK6.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93288394.
RA   SAJJADI F.G., PASQUALE E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
DR   EMBL; Z19110; CAA79526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
FT   NON_TER        1      1
SQ   SEQUENCE    952 AA;   105994 MW;   D939D7D0 CRC32;
```

```
Query Match              45.1%;   Score 541.5;   DB 13;   Length 952;
Best Local Similarity    45.4%;   Pred. No. 7.7e-45;
Matches 103;  Conservative 39;  Mismatches 46;  Indels 39;  Gaps 5;
```

```
Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL   60
        |||:||:|:   ||| ||||||:  ::||||| | :  |  ::|||:|:|||||:||:||:|
Db   23 WEEVSGYDENLNTIRIYQVCNVFEPNQNNWLLTTFINRRGAHRIYTEMRFTVRDCSSLPN   82

Qy   61 VLGTCKEEFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTE  116
        | |:||| ||||||| |||:|     ::|      :|||:|||||||:| | | |::|:
Db   83 VPGSCKEIFNLYYETDSVIATKKSAEWTEAPYLKVDTIAADESFSQVDFGGRLMK----  139

Qy  117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV  175
                                          |  :|||||| :|:|  || :   :||
Db  139 --------------GXEFKKCPSVVQNFAIFPEIMTGAESTSLVTA                170

Qy  176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ  220
        || |:   :|| ||    |: ||:||||||:|:| ||| |  ::|:
Db  171 RGTCIPNAEEVDVPIKLYCNGDGEWMVPIGRCTCKAGYEPENNVACR  217
```

Search completed: May 9, 2000, 22:31:07

This Page Blank (uspto)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 21:52:09 ; Search time 49.21 Seconds
(without alignments)
136.153 Million cell updates/sec

Title: US-09-104-340-1
Perfect score: 1200
Sequence: 1 WEEISGVDEHYTPIRIYQVC............VPIGKCSCNAGYEERGPMCQ 220

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 3045973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1200	100.0	983	1	EPA3_HUMAN	P29320: homo sapien
2	1186	98.8	984	1	EPA3_RAT	O08680 rattus norv
3	1176.5	98.0	983	1	EPA3_MOUSE	P29319 mus musculu
4	1139	94.9	983	1	EPA3_CHICK	P29318 gallus gall
5	914.5	76.2	948	1	EPA6_RAT	P54758 rattus norv
6	914.5	76.2	1035	1	EPA6_MOUSE	Q62413 mus musculu
7	910.5	75.9	986	1	EPA4_CHICK	Q07496 gallus gall
8	905.5	75.5	986	1	EPA4_HUMAN	P54764 homo sapien
9	892.5	74.4	986	1	EPA4_MOUSE	Q03137 mus musculu
10	866.5	72.2	877	1	EPA5_MOUSE	Q60629 mus musculu
11	863.5	72.0	1005	1	EPA5_RAT	P54757 rattus norv
12	860.5	71.7	1037	1	EPA5_HUMAN	P54756 homo sapien
13	857.5	71.5	998	1	EPA7_MOUSE	Q61772 mus musculu
14	856.5	71.4	998	1	EPA7_HUMAN	Q15375 homo sapien
15	856.5	71.4	998	1	EPA7_RAT	P54759 rattus norv
16	831.5	69.3	1004	1	EPA8_MOUSE	O09127 mus musculu
17	825.5	68.8	1013	1	EPA5_CHICK	P54755 gallus gall
18	713	59.4	988	1	EPB2_CHICK	P28693 gallus gall
19	708	59.0	986	1	EPB2_HUMAN	P29323 homo sapien
20	706	58.8	993	1	EPB2_MOUSE	P54763 mus musculu
21	693.5	57.8	984	1	EPB1_RAT	P09759 rattus norv
22	688.5	57.4	984	1	EPB1_HUMAN	P54762 homo sapien
23	681	56.8	987	1	EPB2_COTJA	Q90344 coturnix co
24	637.5	53.1	977	1	EPA2_MOUSE	Q03145 mus musculu
25	632	52.7	998	1	EPA2_HUMAN	P54753 homo sapien
26	627	52.2	976	1	EPA2_HUMAN	P29317 homo sapien
27	620	51.7	993	1	EPB3_MOUSE	P54754 mus musculu
28	574	47.8	1002	1	EPB5_CHICK	Q07497 gallus gall
29	539	44.9	976	1	EPA1_HUMAN	P21709 homo sapien
30	478	39.8	987	1	EPB4_MOUSE	P54761 mus musculu
31	465	38.8	987	1	EPB4_HUMAN	P54760 homo sapien
32	460	38.3	1006	1	EPB6_HUMAN	O15197 homo sapien
33	88.5	7.4	982	1	MSHM_SARGL	O63852 sarcophyton
34	84.5	7.0	3110	1	LMA2_HUMAN	P24043 homo sapien
35	82.5	6.9	461	1	HMCS_ARATH	P54873 arabidopsis
36	80	6.7	819	1	EFG2_YEAST	P39677 saccharomyc
37	79.5	6.6	3106	1	LMA2_MOUSE	Q60675 mus musculu
38	78.5	6.5	158	1	YKP5_KLULA	P05471 kluyveromyc
39	77.5	6.5	1634	1	DPOL_METJA	Q58295 methanococc
40	77	6.4	806	1	SYL_HELPJ	Q92363 helicobacte
41	76.5	6.4	1350	1	MSH6_HUMAN	P52701 homo sapien
42	75.5	6.3	1046	1	CHID_VIBFU	P96156 vibrio furn
43	74.5	6.2	427	1	CG23_YEAST	P24870 saccharomyc
44	74.5	6.2	769	1	ITB2_PIG	P53714 sus scrofa
45	74.5	6.2	771	1	ITB2_MOUSE	P11835 mus musculu

ALIGNMENTS

RESULT 1
EPA3_HUMAN

ID   EPA3_HUMAN   STANDARD;   PRT;   983 AA.
AC   P29320;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (HEK).
GN   EPHA3 OR ETK1 OR ETK OR HEK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92147681.
RA   Boyd A.W., Ward L.D., Wicks I.P., Simpson R.J., Salvaris E., Wilks A.,
RA   Welch K., Loudovaris M., Rockman S., Busmanis I.;
RT   "Isolation and characterization of a novel receptor-type protein
RT   tyrosine kinase (hek) from a human pre-B cell line.";
RL   J. Biol. Chem. 267:3262-3267(1992).
RN   [2]
RP   SEQUENCE OF 21-29 AND 840-860, AND CHARACTERIZATION.
RX   MEDLINE; 92179233.
RA   Wicks I.P., Wilkinson D., Salvaris E., Boyd A.W.;
RT   "Molecular cloning of HEK, the gene encoding a receptor tyrosine
RT   kinase expressed by human lymphoid tumor cell lines.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1611-1615(1992).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A2, -A3, -A4 AND -A5. COULD PLAY A ROLE IN LYMPHOID
CC   FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO LYMPHOID TUMOR CELL LINES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

DR   EMBL; M83941; AAA58633.1; -.
DR   EMBL; A28003; CAA01906.1; -.
DR   PIR; A38224; A38224.
DR   HSSP; P00523; 2PTK.
DR   MIM; 179611; -.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00107; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   20
FT   CHAIN   21   983   EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN   21   541   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   542   565   POTENTIAL.
FT   DOMAIN   566   983   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   21   320   CYS-RICH.
FT   DOMAIN   321   431   FIBRONECTIN TYPE-III.
FT   DOMAIN   432   528   FIBRONECTIN TYPE-III.
FT   DOMAIN   621   882   PROTEIN KINASE.
FT   NP_BIND   627   635   ATP (BY SIMILARITY).
FT   BINDING   653   653   ATP (BY SIMILARITY).
FT   ACT_SITE   746   746   BY SIMILARITY.
FT   MOD_RES   779   779   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD   232   232   POTENTIAL.
FT   CARBOHYD   337   337   POTENTIAL.
FT   CARBOHYD   391   391   POTENTIAL.
FT   CARBOHYD   404   404   POTENTIAL.
FT   CARBOHYD   493   493   POTENTIAL.
FT   CONFLICT   507   507   F -> L (IN CAA01906).
FT   CONFLICT   724   724   V -> L (IN CAA01906).
SQ   SEQUENCE   983 AA; 110086 MW; B8D900FA80FF5121 CRC64;

Query Match   100.0%; Score 1200; DB 1; Length 983;
Best Local Similarity   100.0%; Pred. No. 5.6e-103;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||||||||||||||||||||||||||||||||||||||
Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271

RL   Am. J. Physiol. 274:H331-H341(1998).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; U69278; AAC06273.1; -.
DR   HSSP; P16109; 1FSB.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00536; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   20
FT   CHAIN   21   984   EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN   21   541   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   542   565   POTENTIAL.
FT   DOMAIN   566   984   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   21   321   CYS-RICH.
FT   DOMAIN   322   432   FIBRONECTIN TYPE-III.
FT   DOMAIN   433   529   FIBRONECTIN TYPE-III.
FT   DOMAIN   622   883   PROTEIN KINASE.
FT   NP_BIND   628   636   ATP (BY SIMILARITY).
FT   BINDING   654   654   ATP (BY SIMILARITY).
FT   ACT_SITE   747   747   BY SIMILARITY.
FT   MOD_RES   780   780   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD   232   232   POTENTIAL.
FT   CARBOHYD   337   337   POTENTIAL.
FT   CARBOHYD   391   391   POTENTIAL.
FT   CARBOHYD   404   404   POTENTIAL.
FT   CARBOHYD   493   493   POTENTIAL.
SQ   SEQUENCE   984 AA; 110227 MW; F170C493121F7A0AB CRC64;

Query Match   98.8%; Score 1186; DB 1; Length 984;
Best Local Similarity   98.6%; Pred. No. 1.1e-101;
Matches 217; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          |||||||||||||||||||||||||||||||||||||:||
Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271

RESULT 2
ID   EPA3_RAT   STANDARD;   PRT;   984 AA.
AC   O08680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR REK4).
DE   EPHA3 OR REK4.
GN   EPHA3 OR REK4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE; 98120505.
RA   Li Y.Y., McIernan C.F., Feldman A.M.;
RT   "IL-1 beta alters the expression of the receptor tyrosine kinase gene
RT   r-EphA3 in neonatal rat cardiomyocytes.";

Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQ 271

RESULT 3
EPA3_MOUSE

```
ID   EPA3_MOUSE     STANDARD;     PRT;     983 AA.
AC   P29319;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (MEK4).
GN   EPHA3 OR ETK1 OR MEK4 OR TYRO4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR X SWISS WEBSTER; TISSUE=EMBRYO;
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SECRETED VARIANT WITHOUT THE TRANSMEMBRANE
CC       REGION AND THE KINASE DOMAIN IS PRODUCED BY ALTERNATIVE SPLICING
CC       OF THE SAME GENE.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; M68513; AAA39521.1; -.
DR   EMBL; M68515; AAA39522.1; ALT_SEQ.
DR   PIR; A45583; A45583.
DR   HSSP; P16109; 1FSB.
DR   MGD; MGI:99612; EPHA3.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     20
FT   CHAIN        21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    541    564       POTENTIAL.
FT   DOMAIN      565    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN      432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    231    231       POTENTIAL.
FT   CARBOHYD    336    336       POTENTIAL.
FT   CARBOHYD    390    390       POTENTIAL.
FT   CARBOHYD    403    403       POTENTIAL.
FT   CARBOHYD    492    492       POTENTIAL.
FT   VARSPLIC    530    983       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE    983 AA;    109955 MW;    BE44A6555D8107A2 CRC64;
```

```
Query Match           98.0%;   Score 1176.5;   DB 1;   Length 983;
Best Local Similarity 98.6%;   Pred. No. 8.3e-101;
Matches  217;  Conservative   2;  Mismatches   0;  Indels   1;  Gaps   1;

Qy      1  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     52  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 111

Qy     61  VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    112  VLGTCKETFNLYYMES-DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 170

Qy    121  GPVRKRGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171  GPVNKRGFYILAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Qy    181  NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
           ||||||||||||||||||||| ||||||||||||||| ||
Db    231  NNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQ 270
```

RESULT 4
EPA3_CHICK

```
ID   EPA3_CHICK     STANDARD;     PRT;     983 AA.
AC   P29318;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (CEK4).
GN   EPHA3 OR ETK1 OR CEK4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---

DR   EMBL; M68514; AAA48666.1; -.
DR   PIR; B45583; B45583.
DR   HSSP; P00523; 2PTK.
PRINTS; PR00014; FNTYPEIII.
PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
PFAM; PF00041; fn3; 2.
PFAM; PF00069; pkinase; 1.
PFAM; PF00536; SAM; 1.
PFAM; PF01404; Eph_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   19
FT   CHAIN   20   983   EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN   20   540   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   541   565   POTENTIAL.
FT   DOMAIN   566   983   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   20   320   CYS-RICH.
FT   DOMAIN   321   431   FIBRONECTIN TYPE-III.
FT   DOMAIN   432   528   FIBRONECTIN TYPE-III.
FT   DOMAIN   621   882   PROTEIN KINASE.
FT   NP_BIND   627   635   ATP (BY SIMILARITY).
FT   BINDING   653   653   ATP (BY SIMILARITY).
FT   ACT_SITE   746   746   BY SIMILARITY.
FT   MOD_RES   779   779   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD   231   231   POTENTIAL.
FT   CARBOHYD   336   336   POTENTIAL.
FT   CARBOHYD   390   390   POTENTIAL.
FT   CARBOHYD   403   403   POTENTIAL.
FT   CARBOHYD   492   492   POTENTIAL.
SQ   SEQUENCE   983 AA; 109910 MW; E8895F0BDF77651E CRC64;

Query Match   94.9%; Score 1139; DB 1; Length 983;
Best Local Similarity   94.5%;
Matches 208; Conservative   6; Mismatches 6; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   51 WEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPL 110

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111 VLGTCKETFNLYYMESDDDHLAKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEVREV 170

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GPVSKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          |||||||||||||||||||||||||||||||||||||||:
Db   231 NHSKEEPPKMYCSTEGEWLVPIGKCLCNAGYEERGFPACQ 270

OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 94057777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   HSSP; P00523; 2PTK.
PRINTS; PR00014; FNTYPEIII.
PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
PFAM; PF00041; fn3; 2.
PFAM; PF00069; pkinase; 1.
PFAM; PF01404; Eph_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   22
FT   CHAIN   23   948
FT   DOMAIN   23   549   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   550   570   POTENTIAL.
FT   DOMAIN   571   948   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   630   943   PROTEIN KINASE.
FT   NP_BIND   636   644   ATP (BY SIMILARITY).
FT   BINDING   662   662   ATP (BY SIMILARITY).
FT   ACT_SITE   797   797   BY SIMILARITY.
FT   CARBOHYD   342   342   POTENTIAL.
FT   CARBOHYD   396   396   POTENTIAL.
FT   CARBOHYD   409   409   POTENTIAL.
SQ   SEQUENCE   948 AA; 106235 MW; A47DC78EDB2DEF30 CRC64;

Query Match   76.2%; Score 914.5; DB 1; Length 948;
Best Local Similarity   74.0%;
Matches 159; Conservative 31; Mismatches 24; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          ||||||||||||||||||||||||| || || :|||||||||||||||||||||||||||
Db   56 WDAITEMDEHRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIVEMKFTLRDCNSIPW 115

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||:|: |||| || :|| |:|||||||||||||||||||||||||||:|||
Db   116 VLGTCKETFLYYIESDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREV 175

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          |||||||||||||||:||||||||||||||||||||||||||||||| :|| |||||||
Db   176 GPIERGFYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSC 235

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 214
          |: ||||||||||||||||||||||||| :||||
Db   236 VKSSEERDTPKLYCGADGDWLVPLGRCICTTGYEE 270

DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 6 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-2) (EPH HOMOLOGY KINASE-2).
GN   EPHA6 OR EHK2 OR EHK-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE; 97047913.
RA   Lee A.M., Navaratnam D., Ichimiya S., Greene M.I., Davis J.G.;
RT   "Cloning of m-ehk2 from the murine inner ear, an eph family receptor
RT   tyrosine kinase expressed in the developing and adult cochlea.";
RL   DNA Cell Biol. 15:817-825(1996).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

DR   EMBL; U58332; AAB53836.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:108034; EPHA6.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   22   POTENTIAL.
FT   CHAIN   23   1035   EPHRIN TYPE-A RECEPTOR 6.
FT   DOMAIN   23   549   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   550   570   POTENTIAL.
FT   DOMAIN   571   1035   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   630   943   PROTEIN KINASE.
FT   NP_BIND   636   644   ATP (BY SIMILARITY).
FT   BINDING   662   662   ATP (BY SIMILARITY).
FT   ACT_SITE   797   797   BY SIMILARITY.
FT   CARBOHYD   342   342   POTENTIAL.
FT   CARBOHYD   396   396   POTENTIAL.
FT   CARBOHYD   409   409   POTENTIAL.
SO   SEQUENCE 1035 AA; 116137 MW; 560B2641945A5EF74 CRC64;

Query Match   76.28; Score 914.5; DB 1; Length 1035;
Best Local Similarity 74.0%;
Matches 159; Conservative 33; Mismatches 22; Indels 1; Gaps 1;

```
Qy    1 WEEISGGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL  60
          |:|: :|||  :||  :|||||||||:| |||||||||::|:|||||::||||||||||||
Db   56 WDAITEMDEHNRPIHIYQVCNVMEPNQNNWLRTNWISRDAAQKIYVEMKFTLRDCNSIPW  115

Qy   61 VLGTCKETFNIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV    120
```

RESULT   7
EPA4_CHICK

ID   EPA4_CHICK   STANDARD;   PRT;   986 AA.
AC   Q07496; Q90772;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR CEK8).
GN   EPHA4 OR CEK8.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 93288394.
RA   Sajjadi F.G., Pasquale E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
RN   [2]
RP   SEQUENCE OF 138-986 FROM N.A.
RC   TISSUE=SPINAL CORD;
RX   MEDLINE; 96404128.
RA   Ohta K., Nakamura M., Hirokawa K., Tanaka S., Iwama A., Suda T.,
RA   Ando N., Tanaka H.;
RT   "The receptor tyrosine kinase, Cek8, is transiently expressed on
RT   subtypes of motoneurons in the spinal cord during development.";
RL   Mech. Dev. 54:59-69(1996).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.

DR   EMBL; D38174; BAA07373.1; -.
DR   EMBL; Z19059; CAA79509.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.

```
Db  116 VLGTCKETFNLYYIESDDSHGIFKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEIREV  175
          |||||||||:| |:||| ||:|  |||  |||||||||||||||||||||||||||||||||
Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVFKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC   179
          |: :|||:|||||:|||||:|||||||| |||||:|:||||||:| : |:|||||||||
Db  176 GPIERKGFYLAFQDIGACIALVSVRVFVKCPFTVRSLAMFPDTIPRVDSSSLVEVRGSC  235
Qy  180 VNNSKEDPPRMYCSTEGEWLVPIGKCSCNAGYEE   214
          |:::|| ||::||| ||:||||:| | ||:|||||
Db  236 VKSAEERDTPKLYCGADGDWLVPLGRCICSTGYEE  270
```

RESULT   8

EPA4_HUMAN

ID   EPA4_HUMAN   STANDARD;   PRT;   986 AA.
AC   P54764;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (RECEPTOR PROTEIN-TYROSINE KINASE HEK8)
GN   EPHA4 OR SEK OR HEK8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   -!-   EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   -!-   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; L36645; AAA74246.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602188; -.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; PROTEIN_KINASE; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   19
FT   CHAIN   20   986   EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN   20   547   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM 548   569   POTENTIAL.
FT   DOMAIN   570   986   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   325   435   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   436   532   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   621   882   PROTEIN KINASE.
FT   NP_BIND   627   635   ATP (BY SIMILARITY).
FT   BINDING   653   653   ATP (BY SIMILARITY).
FT   ACT_SITE 746   746   BY SIMILARITY.
FT   MOD_RES   779   779   PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD 235   235   POTENTIAL.
FT   CARBOHYD 340   340   POTENTIAL.
FT   CARBOHYD 408   408   POTENTIAL.
FT   CONFLICT 138   138   R -> G (IN REF. 2).
FT   CONFLICT 487   487   S -> T (IN REF. 2).
SQ   SEQUENCE   986 AA; 109482 MW; BD88C2A5BD840A0F CRC64;

Query Match   75.9%; Score 910.5; DB 1; Length 986;
Best Local Similarity   73.3%; Pred. No. 2.9e-76;
Matches 162; Conservative 30; Mismatches 28; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL   60
          |||:| :|| |||||||||||||||||||||||:|||||:||||||||||||||||::|
Db   54 WEEVSIMDEKNTPIRIYQVCNVMEPSQNNWLRTDWIPREGAQRVYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||:||||||||||:||||| |||| :||:||||||||||||||:||||||||||::||:|
Db   114 VMGTCKETFNLYYYESDNDKERFIRESQFAKIDTIAADESFTQVDIGDRIMKLNTEVRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          ||:||||||||||||||||:||||||||:|||:||:||||:|||| |:||||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFKKCPLTVRNLAQFPDTITGADTSSLVEVRGSC   233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKGCSCNAGYEERGFMCQ   220
          |||||||||||||||||:||||||| ||||||||||||||||
Db   234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGYEERNGECQ   274

RESULT   9

EPA4_MOUSE

ID   EPA4_MOUSE   STANDARD;   PRT;   986 AA.
AC   Q03137;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)

Query Match   75.5%; Score 905.5; DB 1; Length 986;
Best Local Similarity   73.3%; Pred. No. 8.2e-76;
Matches 162; Conservative 30; Mismatches 28; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL   60
          |||:| :|| |||||||||||||||||||||||:|||||:||||||||||||||||::|
Db   54 WEEVSIMDEKNTPIRIYQVCNVMEPSQNNWLRTDWIPREGAQRVYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||:||||||||||:||||| |||| :||:||||||||||||||:||||||||||::||:|
Db   114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          ||:||||||||||||||||:||||||||:|||:||:||||:|||| |:||||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFKKCPLTVRNLAQFPDTITGADTSSLVEVRGSC   233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKGCSCNAGYEERGFMCQ   220
          |||||||||||||||||:||||||| ||||||||||||||||
Db   234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ   274

SQ   SEQUENCE   986 AA; 109859 MW; 0C39C1152EDD46F CRC64;

```
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (MPK-3).
GN   EPHA4 OR SEK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 93096484.
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattei M.-G.,
RA   Chester A., Wilkinson D.G., Charnay P.;
RT   "An Eph-related receptor protein tyrosine kinase gene segmentally
RT   expressed in the developing mouse hindbrain.";
RL   Oncogene 7:2499-2506(1992).
RN   [2]
RP   ERRATUM.
RX   MEDLINE; 93205393.
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattei M.-G.,
RA   Chester A., Wilkinson D.G., Charnay P.;
RL   Oncogene 8:1103-1103(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC       MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PROCESS INVOLVED IN
CC       HINDBRAIN PATTERN FORMATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SHORTER FORM WITH A DELETION OF 49
CC       AMINO ACIDS ALTERING THE CATALYTIC SITE MAY BE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE BRAIN, LOWER LEVELS
CC       IN THE HEART, LUNG AND KIDNEY. IT IS SEGMENTALLY EXPRESSED
CC       IN THE EMBRYONIC BRAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; X65138; CAA46268.1; -.
DR   EMBL; X57241; CAA40517.1; -.
DR   EMBL; S57168; AAB25836.1; -.
DR   HSSP; P00523; 2PTK.
MGD; MGI:98277; EPHA4.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT   SIGNAL        1     19
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CARBOHYD    423    423       POTENTIAL.
FT   VARSPLIC    783    832       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   986 AA;  109801 MW;  D16ADB885668C80E CRC64;
```

```
Query Match             74.4%;  Score 892.5;  DB 1;  Length 986;
Best Local Similarity   72.4%;  Pred. No. 1.3e-74;
Matches  160;  Conservative   30;  Mismatches   30;  Indels    1;  Gaps    1;

Qy    1 WEEISGVDEHYPIRTIQVCNVNVMDHSQNNWLRTNWVPRNSAQRIYIVELKFTLRDCNSIPL 60
          |||:| ||:||||:||||:|||||||||||:| ||:| :|||||:||:|||||||||||:|
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRIDWIITREGAQRVYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDIIAADESFTQMDLGDRILKLNTEIREV 120
          |:|||||:||||||:|:|:| ||: |  || |||||||||:|:|:|||||:||||||:|
Db  114 VMGTCKEFFNLYYESDNDKERRIRESQFGKTDIIAADESFIQVDIGDRIMKLNTEIRDV 173

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDIVP-MDSQSLVEVRGSC 179
          ||::|||||||||||||:|||||||||:|||||:|:||||:||||   |||:|||||||
Db  174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDITGADTSSLVEVRGSC 233

Qy  180 VNNSKEDDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||:|:||:||||  :|||||||||:| |||||:|  ||
Db  234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEEQNGSCQ 274
```

```
RESULT 10
ID   EPA5_MOUSE     STANDARD;      PRT;   877 AA.
AC   Q60629;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (BRAIN-SPECIFIC KINASE)
DE   (CEK-7).
GN   EPHA5 OR EHK1 OR CEK7 OR BSK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 94194581.
RA   Zhou R., Copeland T.D., Kromer L.F., Schulz N.T.;
RT   "Isolation and characterization of Bsk, a growth factor receptor-like
RT   tyrosine kinase associated with the limbic system.";
RL   J. Neurosci. Res. 37:129-143(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
```

```
DR   EMBL; U07357; AAA17038.1; -.
```

```
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:99554; EPHA5.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Transferase; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    877       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN       27    412       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    413    433       POTENTIAL.
FT   DOMAIN      434    877       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      305    399       FIBRONECTIN TYPE-III.
FT   DOMAIN      514    775       PROTEIN KINASE.
FT   NP_BIND     520    528       ATP (BY SIMILARITY).
FT   BINDING     546    546       ATP (BY SIMILARITY).
FT   ACT_SITE    639    639       BY SIMILARITY.
FT   CARBOHYD    266    266       POTENTIAL.
FT   CARBOHYD    301    301       POTENTIAL.
SQ   SEQUENCE   877 AA;  97115 MW;  54AD2C86417B214 CRC64;
```

```
Query Match             72.2%;  Score 866.5;  DB 1;  Length 877;
Best Local Similarity   68.3%;
Matches  151;  Conservative   35;  Pred. No. 2.8e-72;  Mismatches   34;  Indels   1;  Gaps   1;

Qy      1  WEEISGVDEHYPIRTYQVCNVNDHSQNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL  60
Db     85  WEEIGEVDENYAPIHTYQVCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy     61  VLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
Db    145  GLGTCKETFNMYYFESDDENGRSIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy    121  GPVNKKGFYLAFQDVGACAVLVSVRVFKKCPFTVKNLAMFPDVP-MDSQSLVEVRGSC  179
Db    205  GPLSKKGFYLAFQDVGACIALVSVRVIYKKCPSVRVHLAIFPDITGADSSQLLEVSGSC 264

Qy    180  VNNSKEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
Db    265  VNHSVTDDPPKMICSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

```
RESULT 11
EPA5_RAT
ID   EPA5_RAT      STANDARD;      PRT;  1005 AA.
AC   P54757;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1).
GN   EPHA5 OR EKH1 OR EHK-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94067777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 95206467;
RA   Taylor V., Pfarr S., Miescher G.C., Honegger P., Breitschopf H.,
RA   Lassmann H., Steck A.J.;
RT   "Expression and developmental regulation of Ehk-1, a neuronal
RT   Elk-like receptor tyrosine kinase in brain.";
RL   Neuroscience 63:163-178(1994).
```

```
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC       SYSTEM. PREDOMINANTLY EXPRESSED IN NEURONS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; X78689; CAA55357.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1      ?       POTENTIAL.
FT   CHAIN         ?   1005       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN        ?    575       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    576    596       POTENTIAL.
FT   DOMAIN      597   1005       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      677    938       PROTEIN KINASE.
FT   NP_BIND     683    691       ATP (BY SIMILARITY).
FT   BINDING     709    709       ATP (BY SIMILARITY).
FT   ACT_SITE    802    802       BY SIMILARITY.
FT   CARBOHYD     10     20       POTENTIAL.
FT   CARBOHYD    266    266       POTENTIAL.
FT   CARBOHYD    301    301       POTENTIAL.
FT   CARBOHYD    371    371       POTENTIAL.
FT   CARBOHYD    425    425       POTENTIAL.
FT   CARBOHYD    438    438       POTENTIAL.
FT   CARBOHYD    463    463       POTENTIAL.
FT   VARSPLIC    306    358       VCRPGFFKASPHSQTCSKCPPHSYTHEEASTSCVCEKDYFR
FT                                RESDPPTMACTR -> G (IN SPLICED FORMS).
FT   VARSPLIC    358    470       RPPSAPRNAISNVNETSVELEWIPPADTGGGKDVSYILCK
FT                                KCNSHAGVCEECGGHVRYLPQOIGLKNTSVMMADPLAHTNY
FT                                TFEIEAVNGVSDLSPGTRQYVSVNVTTNQAA -> T (IN
FT                                SPLICED FORMS).
FT   VARSPLIC    597    621       SGSCECGCGCRASSLCAVAHPSLIW -> R (IN
FT                                SPLICED FORMS).
FT   CONFLICT    170    170       D -> E (IN REF. 2).
FT   CONFLICT    566    566       G -> A (IN REF. 2).
FT   CONFLICT    578    578       G -> A (IN REF. 2).
```

```
FT   CONFLICT     669    669       G -> A (IN REF. 2).
FT   CONFLICT     708    708       T -> I (IN REF. 2).
FT   CONFLICT     979    979       T -> I (IN REF. 2).
FT   CONFLICT     979    979       T -> I (IN REF. 2).
SQ   SEQUENCE   1005 AA;  111007 MW;  1AED42C99G93C574 CRC64;
```

RESULT 12

```
Query Match    Best Local Similarity   72.0%;   Score 863.5;   DB 1;   Length 1005;
Matches 150;  Conservative 36;  Mismatches 34;  Indels  1;  Gaps  1;

Qy    1  WEEISGVDEHTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         ||||  |||||| ||||  ||||| ||| ||  :||||||||||||||||||||::|
Db   85  WEEIGEVDENYAPIHTYQVCKVMEQNONNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy   61  VLGTCKETFNLIYMESDDDHGVKFREHOFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||:|| |||  ||:|| :|||:|| :||||||::|||| :|||||:||||:|
Db  145  GLGTCKETFNMYYFESDDONGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy  121  GPVNKKGFYLAFQDVGACVALVSVRYVFKKCPFTVKNLAMEFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||| ||||||| |||||||||| |||:|||:||| ||||||||||||
Db  205  GPLSKKGFYLAFQDVGACIALVSVSRVVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 264

Qy  180  VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||:| |||||||||||||||||||:||||||||||| |||||
Db  265  VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEERKNGTCQ 305
```

```
ID   EPA5_HUMAN     STANDARD;      PRT;   1037 AA.
AC   P54756;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (RECEPTOR PROTEIN-
DE   TYROSINE KINASE HEK7).
GN   EPHA5 OR EHK1 OR HEK7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Miescher G.C.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 25-1037 FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC       SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; X95425; CAA64700.1; -.
DR   EMBL; L36644; AAA74245.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 600004; -.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     24
FT   CHAIN        25   1037       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN       25    573       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    574    594       POTENTIAL.
FT   DOMAIN      595   1037       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      675    936       PROTEIN KINASE.
FT   NP_BIND     681    689       ATP (BY SIMILARITY).
FT   BINDING     707    707       ATP (BY SIMILARITY).
FT   ACT_SITE    800    800       BY SIMILARITY.
FT   CARBOHYD    264    264       POTENTIAL.
FT   CARBOHYD    299    299       POTENTIAL.
FT   CARBOHYD    369    369       POTENTIAL.
FT   CARBOHYD    423    423       POTENTIAL.
FT   CARBOHYD    436    436       POTENTIAL.
FT   CARBOHYD    461    461       POTENTIAL.
FT   VARSPLIC    597    619       SCCECGCGRASSLCAVAHPILIW -> R (IN ISOFORM
FT                                2).
SQ   SEQUENCE   1037 AA;  114784 MW;  FC2C46C959AFB699 CRC64;
```

RESULT 13
EPA7_MOUSE

```
Query Match    Best Local Similarity   71.9%;   Score 860.5;   DB 1;   Length 1037;
Matches 150;  Conservative 35;  Mismatches 35;  Indels  1;  Gaps  1;

Qy    1  WEEISGVDEHTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         ||||  |||||| ||||  ||||| ||| ||  :|||||||||||||||||||:|
Db   83  WEEIGEVDENYAPIHTYQVCKVMEQNONNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 142

Qy   61  VLGTCKETFNLIYMESDDDHGVKFREHOFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||:|| |||  ||:|| :|||:|| :||||||::|||| :|||||:||||:|
Db  143  GLGTCKETFNMYYFESDDONGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 202

Qy  121  GPVNKKGFYLAFQDVGACVALVSVRYVFKKCPFTVKNLAMEFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||| ||||||| |||||||||| |||:|||:||| ||||||||||||
Db  203  GPLSKKGFYLAFQDVGACIALVSVSRVVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 262

Qy  180  VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||:| |||||||||||||||||||:||||||||||| |||||
Db  263  VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEERKNGTCQ 303
```

```
ID   EPA7_MOUSE     STANDARD;      PRT;    998 AA.
AC   061772; 061773; 061774; 061505;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3) (EMBRYONIC BRAIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE 1) (MDK-1).
GN   EPHA7 OR EHK3 OR EBK OR MDK1.
```

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 95124729.
RA   Ciossek T., Millauer B., Ullrich A.;
RT   "Identification of alternatively spliced mRNAs encoding variants of
RT   MDK1, a novel receptor tyrosine kinase expressed in the murine
RT   nervous system.";
RL   Oncogene 10:97-108(1995).
RN   [2]
RP   SEQUENCE OF 431-998 FROM N.A. (ISOFORM MDK1).
RC   TISSUE=BRAIN;
RX   MEDLINE; 96081374.
RA   Ellis J., Liu Q., Breitman M., Jenkins N.A., Gilbert D.J.,
RA   Copeland N.G., Tempest H.V., Warren S., Muir E., Schilling H.,
RA   Fletcher F.A., Ziegler S.F., Rogers J.H.;
RT   "Embryo brain kinase: a novel gene of the eph/elk receptor tyrosine
RT   kinase family.";
RL   Mech. Dev. 52:319-341(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: FIVE FORMS OF THE PROTEIN (MDK1, MDK1-1,
CC       MDK1-2, MDK1-T1 AND MDK1-T2) ARE PRODUCED BY ALTERNATIVE
CC       SPLICING OF THE SAME GENE. THE SEQUENCE SHOWN HERE IS THAT OF
CC       MDK1.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN EMBRYO. IN ADULT,
CC       EXPRESSION RESTRICTED TO HIPPOCAMPUS, TESTIS AND SPLEEN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X79082; CAA55687.1; -.
DR   EMBL; X79083; CAA55688.1; -.
DR   EMBL; X79084; CAA55689.1; -.
DR   EMBL; X81466; CAA57224.1; -.
HSSP; P00523; 2PTK.
DR   MGD; MGI:95276; EPHA7.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL          1     29
FT   CHAIN          30    998       EPHRIN TYPE-A RECEPTOR 7.
FT   DOMAIN         30    556       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      557    577       POTENTIAL.
FT   DOMAIN        578    998       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        329    438       FIBRONECTIN TYPE-III.
FT   DOMAIN        439    537       FIBRONECTIN TYPE-III.
FT   DOMAIN        633    894       PROTEIN KINASE.
FT   NP_BIND       639    647       ATP (BY SIMILARITY).
FT   BINDING       665    665       ATP (BY SIMILARITY).
FT   ACT_SITE      758    758       BY SIMILARITY.
FT   MOD_RES       791    791       PHOSPHORYLATION (BY SIMILARITY).
FT   CARBOHYD       64     64       POTENTIAL.
FT   CARBOHYD      343    343       POTENTIAL.
FT   CARBOHYD      410    410       POTENTIAL.
FT   VARSPLIC      540    544       MISSING (IN ISOFORM MDK1-1).
FT   VARSPLIC      601    604       MISSING (IN ISOFORM MDK1-2).
FT   VARSPLIC      600    610       FKPFGTKYID -> SLVTNEHLSVL (IN ISOFORM
FT                                  MDK1-T1).
FT   VARSPLIC      611    998       MISSING (IN ISOFORM MDK1-T1).
FT   VARSPLIC      600    626       EKTQHNKKWMIASCSRL (IN ISOFORM MDK1-T2).
FT   CONFLICT      627    998       MISSING (IN ISOFORM MDK1-T2).
FT   CONFLICT      480    480       Y -> H (IN REF. 2).
SQ   SEQUENCE   998 AA;  111874 MW;  81C5538E15AEA2FA CRC64;
```

```
Query Match    71.5%;  Score 857.5; DB 1;  Length 998;
Best Local Similarity   70.0%;  Pred. No. 2.2e-71;
Matches  156;  Conservative  36;  Mismatches  28;  Indels  3;  Gaps  2;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQONNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||||||:||||||||||||||||||| |: ||||||||| ||||||||||||||||||||
Db   55 WEEISGDLDENYTPIRTYQVCQVWEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHOFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||:|||:||:||||| || :|| |||| ||||||||| ||||||| |::|||:||||
Db  115 VLGTCKETFNLYYIETDYDTGRNIRENLYVKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy  121 GPVNKKGFYLAFQDVGACVALYSVRVYFKKCPTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         |:||||||||||||:|||||||:||||:|||||||||||||||||||:||||||||||||:
Db  175 GPLSKKGFYLAFQDVGACIALVSVKVYKKCWIVENLAVFPDTVTGSSEFSLVEVRGTC 234

Qy  180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||:::|  : |||||||||| :|||||||||||||||||:||:
Db  235 VSSAEEEAENSPRMHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277
```

```
RESULT 14
ID   EPA7_HUMAN     STANDARD;      PRT;   998 AA.
AC   Q15375;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3) (RECEPTOR PROTEIN-
DE   TYROSINE KINASE EHK3 OR HEK11).
GN   EPHA7 OR EHK3 OR HEK11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

Left column

```
RESULT 15

ID   EPA7_RAT     STANDARD;      PRT;   998 AA.
AC   P54759;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3).
GN   EPHA7 OR EHK3 OR EHK-3.
OS   Rattus norvegicus (Rat).
```

```
CC   ------------------------------------------------------------------------
DR   EMBL; L36642; AAA74243.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602190; -.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Transferase; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    998       EPHRIN TYPE-A RECEPTOR 7.
FT   DOMAIN       25    556       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    557    577       POTENTIAL.
FT   DOMAIN      578    998       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      329    438       FIBRONECTIN TYPE-III.
FT   DOMAIN      439    537       FIBRONECTIN TYPE-III.
FT   DOMAIN      633    894       PROTEIN KINASE.
FT   NP_BIND     639    647       ATP (BY SIMILARITY).
FT   BINDING     665    665       ATP (BY SIMILARITY).
FT   ACT_SITE    758    758       BY SIMILARITY.
FT   MOD_RES     791    791       PHOSPHORYLATION BY SIMILARITY.
FT   CARBOHYD    343    343       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
FT   MOD_RES     791    791       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   998 AA;  112096 MW;  47989CA0D2BB06EB CRC64;

Query Match             71.4%;  Score 856.5;  DB 1;  Length 998;
Best Local Similarity   69.5%;  Pred. No. 2.8e-71;
Matches 155;  Conservative 37;  Mismatches 28;  Indels  3;  Gaps  2;

Qy      1  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
           |||||:||:|:||||||||||||||||| ||: ||||||| :|||:||||||||||||||:
Db      1  WEEISGLDENVTPIRTYQVCQVMEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy     61  VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
           |||||||||||:||||: |:  |   ||:||||||||||||||: :||:||||||||||:|
Db    115  VLGTCKETFNLYYIETDYDTGRNIRENLYVKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy    121  GPVNKRGFYLAFQDVGACVALVSRVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
           ||:||||||||||||| |||:::| ||||||| | :||:|:|| :||  |||:|||||| |
Db    175  GPLSKKGFYLAFQDVGACIALVSVKVYKKCWSIIENLAIPPDTVTGSEFSSLVEVRGTC 234

Qy    180  VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
           |:|:::| | |||||||||||||||||||||:| |||||||||
Db    235  VSSAEEEAENAPRMHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277
```

Right column

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95249272.
RA   Valenzuela D.M., Rojas E., Griffiths J.A., Compton D.L., Gisser M.,
RA   Ip N.Y., Goldfarb M., Yancopoulos G.D.;
RT   "Identification of full-length and truncated forms of Ehk-3, a novel
RT   member of the Eph receptor tyrosine kinase family.";
RL   Oncogene 10:1573-1580(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE
CC       SPLICING. THE TRUNCATED FORM LACKS THE KINASE DOMAIN. THE LONG
CC       FORM IS MORE WIDELY EXPRESSED IN THE EMBRYO.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO THE NERVOUS SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   ------------------------------------------------------------------------
```

```
CC   ------------------------------------------------------------------------
DR   EMBL; U21954; AAA86830.1; -.
DR   EMBL; U21955; AAA86831.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Transferase; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    998       EPHRIN TYPE-A RECEPTOR 7.
FT   DOMAIN       25    556       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    557    577       POTENTIAL.
FT   DOMAIN      578    998       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      329    438       FIBRONECTIN TYPE-III.
FT   DOMAIN      439    537       FIBRONECTIN TYPE-III.
FT   DOMAIN      633    894       PROTEIN KINASE.
FT   NP_BIND     639    647       ATP (BY SIMILARITY).
FT   BINDING     665    665       ATP (BY SIMILARITY).
FT   ACT_SITE    758    758       BY SIMILARITY.
FT   MOD_RES     791    791       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    343    343       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
FT   VARSPLIC    600    610       FKPGTKIYID -> SLVTNEHLSVL (IN TRUNCATED
FT                                ISOFORM).
SQ   SEQUENCE   998 AA;  111953 MW;  A7A82A69892487GC CRC64;

Query Match             71.4%;  Score 856.5;  DB 1;  Length 998;
Best Local Similarity   69.5%;  Pred. No. 2.8e-71;
Matches 155;  Conservative 37;  Mismatches 28;  Indels  3;  Gaps  2;

Qy      1  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
           |||||:||:|:||||||||||||||||| ||: ||||||| :|||:||||||||||||||:1
```

```
Db   55  WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG  114

Qy   61  VLGTCKETFNLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV  120
         ||||||||||||| |:  | |: |||||||||||||||||:| ||||:|:-|:
Db  115  VLGTCKETFNLIYYETDIDYDIGRNIRENLYVKIDTIAADESFTQGDLGERKMKLNTEVREI  174

Qy  121  GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC  179
         ||:|||||||||||||||:|||||||||||:|||||:|||||  :  |||||||:|
Db  175  GPLSKKGFYLAFQDVGACIALVSVKYYYKKCWSIIENLAVFPDTVTGSEFSSLVEVRGTC  234

Qy  180  VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  220
         |:::::|  |: |||:||||||||||||||||| | |::::|  |:
Db  235  VSSAEEAENSPRMHCSAEGEWLVPIGKCICKAGYQQKGDTCE  277
```

Search completed: May 9, 2000, 22:32:06
Job time: 2397 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 21:18:01 ; Search time 64.29 Seconds
(without alignments)
200.631 Million cell updates/sec

Title:
Perfect score: 1200
Sequence: 1 WEEISGVDEHYTPIRTYQVC..........VPIGKCSCNAGYEBERGFMCQ 220

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_63:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1200	100.0	983	2	A38224	protein-tyrosine k
2	1176.5	98.0	983	2	A45583	receptor tyrosine
3	1139	94.9	983	2	B45583	receptor tyrosine k
4	914.5	76.2	948	2	S51605	receptor-like tyro
5	905.5	75.5	986	2	I78844	receptor protein-t
6	894.5	74.5	985	2	I51549	receptor tyrosine
7	892.5	74.4	986	2	S78059	protein-tyrosine k
8	866.5	72.2	877	2	I48967	brain-specific kin
9	866.5	72.2	898	2	S47489	receptor tyrosine
10	863.5	72.0	893	2	S51603	receptor tyrosine
11	863.5	72.0	981	2	S51604	receptor-like tyro
12	863.5	72.0	991	2	S49015	receptor tyrosine
13	860.5	71.7	991	2	I78843	receptor tyrosine
14	857.5	71.5	605	2	JC5673	receptor tyrosine
15	857.5	71.5	610	2	I48612	developmental kina
16	857.5	71.5	626	2	I48614	developmental kina
17	857.5	71.5	998	2	JC5672	receptor tyrosine
18	856.5	71.4	998	2	I58351	receptor protein-t
19	825.5	68.8	1013	2	I50615	receptor-type prot
20	713	59.4	995	2	A56599	embryo kinase 5 -
21	708	59.0	970	2	I78842	protein protein-t
22	693.5	57.8	984	2	A39753	receptor tyrosine k
23	691.5	57.6	985	2	I51672	receptor tyrosine
24	644	53.7	998	2	I50611	protein-tyrosine k
25	637.5	53.1	977	2	S49004	tyrosine kinase Mp
26	632	52.7	998	2	S37627	protein-tyrosine k
27	627	52.2	976	2	A36355	protein-tyrosine k
28	620	51.7	993	2	I48653	mouse developmenta
29	601	50.1	975	2	I48974	receptor-protein t
30	589	49.1	938	2	I49071	protein kinase - m
31	569.5	47.5	849	2	I50617	protein-tyrosine k
32	541.5	45.1	952	2	I50612	protein-tyrosine k
33	539	44.9	984	1	A34076	protein-tyrosine k
34	478	39.8	984	2	I48953	eph-related recept
35	478	39.8	987	2	I48652	mouse developmenta
36	466	38.8	987	2	A54092	protein-tyrosine k
37	460	38.3	1006	2	JC5526	kinase-defective E
38	369.5	30.8	1019	2	T13039	tyrosine kinase re
39	241	20.1	1122	2	T42400	Eph receptor tyros
40	89.5	7.5	1122	2	T05803	hypothetical prote
41	87.5	7.3	788	2	I51530	integrin beta-3 su
42	86	7.2	919	2	T29581	hypothetical prote
43	85.5	7.1	347	2	T34131	hypothetical prote
44	84.5	7.0	865	2	H71447	trehalose-6-phosph
45	82.5	6.9	369	2	B70220	conserved hypothet

ALIGNMENTS

RESULT 1
A38224
protein-tyrosine kinase (EC 2.7.1.112) hek precursor - human
C;Species: Homo sapiens (man)
A;Accession: A38224
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
R;Wicks, I.P.; Wilkinson, D.; Salvaris, E.; Boyd, A.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 1611-1615, 1992
A;Title: Molecular cloning of HEK, the gene encoding a receptor tyrosine kinase expre
A;Reference number: A38224; MUID:92179233
A;Cross-references: GB:M83941; NID:g183931; PIDN:AAA58633.1; PID:g183932
A;Experimental source: pre-B-cell leukemia cell line LK63
A;Note: sequence extracted from NCBI backbone (NCBIP:86627)
A;Molecule type: mRNA
A;Residues: 1-983 <WIC>
A;Accession: A38224
A;Molecule type: protein
A;Residues: 21-39;810-860 <WI2>
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferas
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-983/Product: protein-tyrosine kinase hek #status experimental <MAT>
F;542-565/Domain: transmembrane #status predicted <TMM>
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;232,337,391,404,493/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 1200; DB 2; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.3e-101;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNLRTNNVPRNSAQKIYVELKFTLRDCNSIPL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNLRTNNVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy  61 VLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRIIKLNTEIREV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRIIKLNTEIREV 171

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTIVPMDSQSLVEVRGSCV 231

Qy 181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
        ||||||||||||||||||||||||||||||||||||||||
Db 232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 2
A45583
receptor tyrosine kinase Mek4 - mouse

C;Species: Mus musculus (house mouse)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse an
A;Reference number: A45583; MUID:92031278
A;Accession: A45583
A;Molecule type: mRNA
A;Status: preliminary
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68513; PIDN:AAA39521.1; PID:g199120
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Note: sequence extracted from NCBI backbone (NCBIN:63398, NCBIP:62401)
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   98.0%; Score 1176.5; DB 2; Length 983;
Best Local Similarity   98.6%; Pred. No. 3.2e-99;
Matches 217; Conservative   2; Mismatches   0; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYYMES-DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 170

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||||||||||||||||||||||||||||||||||||||
Db   231 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 270

RESULT 3
B45583
receptor tyrosine kinase Cek4 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: B45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse an
A;Reference number: A45583; MUID:92031278
A;Accession: B45583
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68514; NID:g454809; PIDN:AAA48666.1; PID:g211447
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Note: sequence extracted from NCBI backbone (NCBIN:62405, NCBIP:62411)
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   94.9%; Score 1139; DB 2; Length 983;
Best Local Similarity   94.5%; Pred. No. 8.4e-96;
Matches 208; Conservative   6; Mismatches   6; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   51 WEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPL 110

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db   111 VLGTCKETFNLYYMESDDDHLAKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEVREV 170

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GPVSKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          |||||||||:||||||||||||||||:||||||||||:||
Db   231 NHSKEEDPPKMYCSTEGEWLVPIGKCLCNAGYEERGFACQ 270

RESULT 4
S51605
receptor-like tyrosine kinase Ehk-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51605
A;Reference number: S49015; MUID:94067777
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-948 <MAI>
A;Cross-references: EMBL:S68030
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;628-936/Domain: protein kinase homology <KIN>
F;636-644/Region: protein kinase ATP-binding motif
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase

Query Match   76.2%; Score 914.5; DB 2; Length 948;
Best Local Similarity   74.0%; Pred. No. 2.3e-75;
Matches 159; Conservative 31; Mismatches 24; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMEPNQNNWLRTNWISRDAAQKIVEMKFTLRDCNSIPW 60
          |:|||||:|||||||||||:||:|||:|||||||| | | |:|||:|||||||||:||:
Db   56 WDAITEMDEHNRPIHITYQVCNVMEPNQNNWLRTNWISRDAAQKIVEMKFTLRDCNSIPW 115

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||:|| :|||:|| : ||||||||||||||||||||||||||||||
Db   116 VLGTCKETFNLYYIESDESHGTKFRPSQYIKIDTIAADESFTQMDLGDRILKLNTEIREV 175

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          ||::|||||||||| |||:||||||||:||||||||:|||||||:| ||:||||||||
Db   176 GPIERKGFYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSC 235

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 214
          |:||:||||||||||||||||||||||:|||||||
Db   236 VKSSEERDTPKLYCGADGDWLVPLGRICTTGYEE 270

RESULT 5
I78844
receptor protein-tyrosine kinase - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999
C;Accession: I78844
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welch
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein
A;Reference number: I58351; MUID:95206782
A;Accession: I78844
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-986 <RES>
A;Cross-references: GB:L36645; NID:g551613; PIDN:AAA74246.1; PID:g551614
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Gene: HEK8
A;Genetics:
F;619-885/Domain: protein kinase homology <KIN>

Query Match   75.5%;   Score 905.5;   DB 2;   Length 986;
Best Local Similarity   73.3%;   Pred. No. 1.6e-74;
Matches 162;   Conservative 30;   Mismatches 28;   Indels   1;   Gaps   1;

```
Qy     1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||:| :|| |||||||||||||||||||| :|| ||||||||||:||||||||||||||:|
Db    54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWIIREGAQRVYIEIKFTLRDCNSLPG 113

Qy    61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         :|||||||||||:||||| :||| |||| |||||||:||||||||||||||||||||:|:|
Db   114 VMGTCKETFNLYYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||||||:||||||||:|| |||||||:||||||:|||||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTIIGADISSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |||:|| :||| |||||||||||||||||:||||||||||||
Db   234 VNNSEEKDVPRMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274
```

A:Reference number: 6
A:Molecule type: S78059
A:Accession: S78059
A:Residues: 1-986 <CHA>
R:Gllardi-Hebenstreit, P.; Nieto, M.A.; Fröhn, M.; Mattei, M.G.; Chestier, A.; Wilkin
Oncogene 7, 2499-2506, 1992
A:Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed 1
A:Reference number: 6
A:Molecule type: mRNA
A:Accession: S30505
A:Molecule type: mRNA
A:Residues: 1-31,55-986 <GIL>
A:Cross-references: EMBL:X65138
C:Genetics:
A:Gene: Sek
C:Genetics:
C:Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransfera
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repe
F:16-986/Product: protein-tyrosine kinase #status predicted <SIG>
F:1-15/Domain: signal sequence #status predicted <SIG>
F:548-569/Domain: transmembrane #status predicted <TMM>
F:619-885/Domain: protein kinase homology <KIN>
F:627-635/Region: protein kinase ATP-binding motif
F:908-974/Domain: SAM homology <SAM>
F:235,340,408,423/Binding site: carbohydrate (Asn) (covalent) #status predicted

---

RESULT 6
I51549
receptor tyrosine kinase - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996
C:Accession: I51549
R:Wining, R.S.; Sargent, T.D.
Mech. Dev. 46, 219-229, 1994
A:Title: Pagliaccio, a member of the Eph family of receptor tyrosine kinase genes, has 1
A:Reference number: I51549; MUID:95001564
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-985 <WIN>
A:Cross-references: GB:L26099; NID:g416402; PIDN:AAA64464.1; PID:g416403
A:Genetics:
A:Gene: Pag
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C:Keywords: ATP; transmembrane protein
F:618-884/Domain: protein kinase homology <KIN>
F:626-634/Region: protein kinase ATP-binding motif

Query Match   74.5%;   Score 894.5;   DB 2;   Length 985;
Best Local Similarity   71.9%;   Pred. No. 1.6e-73;
Matches 159;   Conservative 33;   Mismatches 28;   Indels   1;   Gaps   1;

```
Qy     1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||:| :|| |||||||||||||||||||| :|| ||||||||||:||||||||||||||:|
Db    54 WEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRTDWIITREGAQRVYIEIKFTLRDCNSLPG 113

Qy    61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         :|||||||||||:||||| :||| |||| |||||||:||||||||||||||||||||:|:|
Db   114 VMGTCKETFNLYYYESDNDKERFIRESQFGKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||||||:||||||||:|| |||||||:||||||:|||||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTIIGADISSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |||:|| :||| |||||||||||||||||:||||||||||||
Db   234 VNNSEEKDVPRMYCGADGEWLVPIGNCLCNAGHEEQNGECQ 274
```

A:Reference number: S78059
A:Molecule type: mRNA
A:Accession: S78059
A:Residues: 1-986 <CHA>
R:Gllardi-Hebenstreit, P.; Nieto, M.A.; Fröhn, M.; Mattei, M.G.; Chestier, A.; Wilkin
Oncogene 7, 2499-2506, 1992
A:Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed 1
A:Reference number: S30496; MUID:93096484
A:Molecule type: mRNA
A:Accession: S30505
A:Cross-references: EMBL:X65138

RESULT 7
S78059
protein-tyrosine kinase (EC 2.7.1.112) Eph precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999
C:Accession: S78059; S30505; I58366
R:Charnay, P.
submitted to the EMBL Data Library, March 1992

RESULT 8
I48967
brain-specific kinase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: I48967
R:Zhou, R.R.; Copeland, T.D.; Kromer, L.F.; Schulz, N.T.
J. Neurosci. Res. 37, 129-143, 1994
A:Title: Isolation and characterization of Bsk, a growth factor receptor-like tyrosin
A:Reference number: I48967; MUID:94194581
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-877 <RES>
A:Cross-references: EMBL:U07357; NID:g465369; PIDN:AAA17038.1; PID:g466370
A:Genetics:
A:Gene: Bsk
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C:Keywords: ATP; transmembrane protein
F:512-778/Domain: protein kinase homology <KIN>
F:520-528/Region: protein kinase ATP-binding motif
F:801-868/Domain: SAM homology <SAM>

Query Match   74.4%;   Score 892.5;   DB 2;   Length 986;
Best Local Similarity   72.4%;   Pred. No. 2.5e-73;
Matches 160;   Conservative 30;   Mismatches 30;   Indels   1;   Gaps   1;

```
Qy     1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||:| :|| |||||||||||||||||||| :|| ||||||||||:||||||||||||||:|
Db    54 WEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRTDWIITREGAQRVYIEIKFTLRDCNSLPG 113

Qy    61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         || |||||||||||:| |||| |||| |||||||:||||||||||||||||||||:|:|
Db   114 VMGTCKETFNLYYYESNNDKERFIRESQFGKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||||||:||||||||:|| |||||||:||||||:|||||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTIIGADISSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |||:|| :||| |||||||||||||||||:||||||||||||
Db   234 VNNSEEKDVPRMYCGADGEWLVPIGNCLCNAGHEEQNGECQ 274
```

Query Match   72.2%;   Score 866.5;   DB 2;   Length 877;
Best Local Similarity   68.3%;   Pred. No. 5e-71;

Matches 151; Conservative 35; Mismatches 34; Indels 1; Gaps 1;

```
Qy    1 WEEISGVDEHYPIRTYQVCNVNDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
        |||| :||| ||| |||||||: ::   ||||| :|::   |: |:||||||||||||: :|:
Db   85 WEEIGEVDENYAPIHTYQVCKVKVMEQNQNNWLLISWISNEGASRIFIELKFTLRDCNSLPG 144

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
        ||||||||||: ||:|||| |::|::|: |:||||||||||||:|||||: ||||| :|:
Db  145 GLGTCKETFNMYYFESDDENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
        ||: ||||||||||||||:|||||||||:||| |:|:||:||||: | :||| ||:| |||
Db  205 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 264

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
        ||:| ::|||:||||:||||||||||| | ||||||: |||
Db  265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 9
S47489
receptor tyrosine kinase - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C;Accession: S47489
R;Taylor, V.; Pfarr, S.; Mescher, G.C.; Honegger, P.; Breitschopf, H.; Lassmann, H.; St
submitted to the EMBL Data Library, April 1994
A;Description: Expression and developmental regulation of EHK-1, a neuronal ELK-like rec
A;Reference number: S47489
A;Accession: S47489
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-898 <TAY>
A;Cross-references: EMBL:X78689; NID:g531543; PIDN:CAA55357.1; PID:g531544
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; transmembrane protein
F;568-834/Domain: protein kinase homology <KIN>
F;576-584/Region: protein kinase ATP-binding motif

Query Match   72.2%; Score 866.5; DB 2; Length 898;
Best Local Similarity   68.3%; Pred. No. 5.e-71;
Matches 151; Conservative 35; Mismatches 34; Indels 1; Gaps 1;

```
Qy    1 WEEISGVDEHYPIRTYQVCNVNDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
Db   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
Db  120 GLGTCKETFNMYYFESDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 179

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
Db  179 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 209

Qy  150 GLGTCKETFNMYYFESDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 209

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
Db  210 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 269

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
Db  270 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 310
```

RESULT 10
S51603
receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S51603
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase fam
A;Reference number: S49015; MUID:94067777
A;Accession: S51603
A;Status: preliminary

A;Molecule type: mRNA
A;Residues: 1-893 <MAI>
A;Cross-references: EMBL:S68028
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;563-829/Domain: protein kinase homology <KIN>
F;571-579/Region: protein kinase ATP-binding motif

Query Match   72.0%; Score 863.5; DB 2; Length 893;
Best Local Similarity   67.9%; Pred. No. 9.5e-71;
Matches 150; Conservative 36; Mismatches 34; Indels 1; Gaps 1;

```
Qy    1 WEEISGVDEHYPIRTYQVCNVNDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
Db   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
Db  145 GLGTCKETFNMYYFESDDENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
Db  205 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 264

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
Db  265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 11
S51604
receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51604
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:94067777
A;Accession: S51604
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-981 <MAI>
A;Cross-references: EMBL:S68029
A;Note: the authors translated the codon GAC for residue 170 as Glu
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;651-917/Domain: protein kinase homology <KIN>
F;659-667/Region: protein kinase ATP-binding motif

Query Match   72.0%; Score 863.5; DB 2; Length 981;
Best Local Similarity   67.9%; Pred. No. 1.1e-70;
Matches 150; Conservative 36; Mismatches 34; Indels 1; Gaps 1;

```
Qy    1 WEEISGVDEHYPIRTYQVCNVNDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
Db   85 WEEIGEVDENYAPIHTYQVCKVKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
Db  145 GLGTCKETFNMYYFESDDENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
Db  205 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSC 264

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
Db  265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 12
S49015
receptor tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S49015; S51602
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase fam
A;Reference number: S49015; MUID:94067777
A;Accession: S49015
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1005 <MA1>
A;Cross-references: EMBL:S68024
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Accession: S51602
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-305,'G',359-1005 <MA2>
A;Cross-references: EMBL:S68026
A;Note: the authors translated the codon GAC for residue 170 as Glu
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; transmembrane protein
F;675-941/Domain: protein kinase homology <KIN>
F;683-691/Region: protein kinase ATP-binding motif

Query Match   72.0%; Score 863.5; DB 2; Length 1005;
Best Local Similarity 67.9%; Pred. No. 1.1e-70;
Matches 150; Conservative 36; Mismatches 34; Indels 1; Gaps 1;

```
Qy      1 WEEISGVDEHYTPIRTYQVCNVNDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
            |||| ||||||:||||  ||||| |||:|  |:|||||||||||||||||||||||||||:
Db     85 WEEIGEVDENYAPIHTYQVCKVRWBEQONQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy     61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDIGDRILKLNTEIREV 120
            ||||||||||||||:::||  ||||  |:| |||||||||:|:||||||||||||||:|
Db    145 GLGTCKETFNMYIFESDDENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy    121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDDTVP-MDSQSLVEVRGSC 179
            ||||||||||||||||||| :|||||||||||||||||||||||:|||:|||||||||
Db    205 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTIGADSSQLLEVSGSC 264

Qy    180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
            |:| ||||||||||||||| :|||||||||| :|||||||
Db    265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 13
I78843
receptor protein-tyrosine kinase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999
C;Accession: I78843
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welcher,
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein-ty
A;Reference number: I58351; MUID:95206782
A;Accession: I78843
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-991 <RES>
A;Cross-references: GB:L36644; NID:g551611; PIDN:AAA74245.1; PID:g551612
A;Gene: HEK7
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
F;627-893/Domain: protein kinase homology <KIN>
F;916-982/Domain: SAM homology <SAM>

Query Match   71.8%; Score 860.5; DB 2; Length 991;

Best Local Similarity 67.9%; Pred. No. 2e-70;
Matches 150; Conservative 35; Mismatches 35; Indels 1; Gaps 1;

```
Qy      1 WEEISGVDEHYTPIRTYQVCNVNDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
            |||| |||||:||||  ||||| |||:|  |:||||||||||||||||||||||||||:
Db     59 WEEIGEVDENYAPIHTYQVCKVRWBEQONQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 118

Qy     61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDIGDRILKLNTEIREV 120
            ||||||||||||||:::||  ||||  |:| |||||||||:|:||||||||||||||:|
Db    119 GLGTCKETFNMYIFESDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 178

Qy    121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDDTVP-MDSQSLVEVRGSC 179
            ||||||||||||||||||| :||||||||||||||||||||||:|||:|||||||||
Db    179 GPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAVFPDTIGADSSQLLEVSGSC 238

Qy    180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
            |:| ||||||||||||||| :|||||||||| :|||||||
Db    239 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 279
```

RESULT 14
JC5673
receptor tyrosine kinase (EC 2.7.-.-) Ebk-td1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change 07-May-1999
C;Accession: JC5673
R;Taludker, A.H.; Muramatsu, T.; Kaneda, N.
Cell Struct. Funct. 22, 477-485, 1997
A;Title: A novel truncated variant form of Ebk/MDK1 receptor tyrosine kinase is expre
A;Reference number: JC5672; MUID:98035156
A;Accession: JC5673
A;Molecule type: mRNA
A;Residues: 1-605 <TAL>
A;Experimental source: embryo
C;Comment: This enzyme plays a regulatory role during neural development and embryo
C;Keywords: brain; phosphotransferase
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-605/Product: receptor tyrosine kinase Ebk-td1 #status predicted <MAT>
F;31-548/Domain: extracellular #status predicted <EXT>
F;331-438/Domain: fibronectin type III repeat <FN1>
F;441-534/Domain: fibronectin type III repeat <FN2>
F;549-574/Domain: transmembrane #status predicted <TMM>

Query Match   71.5%; Score 857.5; DB 2; Length 605;
Best Local Similarity 70.0%; Pred. No. 2.1e-70;
Matches 156; Conservative 36; Mismatches 28; Indels 3; Gaps 2;

```
Qy      1 WEEISGVDEHYTPIRTYQVCNVNDHSQNNWLRINWVPRNSAQKIYVELKFTLRDCNSIPL 60
            |||| ||||||:||||  ||||| |||:|  ||||||||||||||||||||||||||||:
Db     55 WEEISGLIDENYTPIRTYQVCQVMEPNQNNWLRINWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy     61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDIGDRILKLNTEIREV 120
            |||||||| |||| |:||  |||  |||||||||||||||| |||:| |:||||||||||:
Db    115 VLGTCKETFNLYYETDIDTGRNIRENLYVKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy    121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDDTVP-MDSQSLVEVRGSC 179
            ||||||||||||||||||| |||||||:|||||||||||||| |||:|||||||||
Db    175 GPLSKKGFYLAFQDVGACIALVSVRVYKKCWTIVENLAVFPDDTVGSEFSSLVEVRGTC 234

Qy    180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
            |:|    |||||||||||||||| :|||||||| :|||||||
Db    235 VSSAEEEAENSPRMHCSAEGEWLVPIGKCICKAGYQKGDTCE 277
```

RESULT 15
I48612
developmental kinase 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48612; S51741
R;Ciossek, T.; Millauer, B.; Ullrich, A.
Oncogene 9, 97-108, 1995

A;Title: Identification of alternatively spliced mRNAs encoding variants of MDK1, a nove
A;Reference number: I48611
A;Accession: I48612
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-610 <RES>
A;Cross-references: EMBL:X79083; NID:g607135; PIDN:CAA55688.1; PID:g607136

Query Match   71.5%;   Score 857.5;   DB 2;   Length 610;
Best Local Similarity   70.0%;   Pred. No. 2.1e-70;
Matches 156; Conservative   36; Mismatches   28;   Indels   3;   Gaps   2;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         |||||||:|:|||||||||||||| || :|||||||| : :||:||:|||||||||||||:|
Db   55 WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||| ||| | |:|| | |||:|||||||||:|| ||||||||| ||||||||||||:|
Db 115 VLGTCKETFNLYYVETDYDTGRNIRENLYVKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:||||||||||||||||||||||:||:|||||||:|||||| :|||||||||||||:
Db 175 GPLSKKGFYLAFQDVGACIALVSVRVYKKCWTIVENLAVFPDTVTGSEFSSLVEVRGTC 234

Qy 180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |:::|| |::|| |||||||| | |||::|| :|
Db 235 VSSAEEAENSPRMHCSAEGEWLVPIGKCICKAGYQQKGDTCE 277

Search completed: May 9, 2000, 22:27:14
Job time: 4153 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 20:49:28 ; Search time 50.71 Seconds
                                    (without alignments)
                                    102.760 Million cell updates/sec

Title:
Perfect score: 1200
Sequence: 1 WEEISGVDEHYTPIRTYQVC.........VPIGKCSCNAGYEERGFMCQ 220

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1200	100.0	983	1	R31466	HEK polypeptide. R
2	1136	94.7	983	1	R75711	Eph-related PTK Ce
3	914.5	76.2	948	1	W83148	Rat receptor tyros
4	905.5	75.5	986	1	W85936	Protein tyrosine-k
5	905.5	75.5	986	1	W85091	EPH-like receptor
6	877	72.2	1005	1	W71628	Mouse Bsk receptor
7	866.5	72.2	877	1	W83147	Rat receptor tyros
8	865.5	72.1	928	1	R97853	Rat REK7 eph-relat
9	860.5	71.7	991	1	R85090	EPH-like receptor
10	860.5	71.5	610	1	R85422	Mouse developmenta
11	857.5	71.5	626	1	W03423	Mouse developmenta
12	857.5	71.5	998	1	W03421	Mouse developmenta
13	856.5	71.4	998	1	W85092	EPH-like receptor
14	713	59.4	995	1	R75712	Eph-related PTK Ce
15	713	59.4	1011	1	R75709	Eph-related PTK Ce
16	708	59.0	970	1	R85089	EPH-like receptor
17	706	58.8	994	1	R87018	Receptor tyrosine
18	706	58.8	994	1	W26366	Mouse Nuk tyrosine
19	694.5	57.9	994	1	R44513	elk. Expression of
20	681.5	56.8	710	1	R75714	Eph-related PTK Ce
21	681.5	56.8	722	1	R75705	Eph-related PTK Ce
22	681.5	56.8	744	1	R75713	Eph-related PTK Ce
23	644	53.7	973	1	R75708	Eph-related PTK Ce
24	644	53.7	988	1	R75710	Eph-related PTK Ce
25	632	52.7	988	1	R51999	Human embryonal ki
26	626	52.2	993	1	R75843	Protein p140 cDNA
27	626	52.2	993	1	R75844	Protein p140 cDNA
28	574	47.8	973	1	R75707	Eph-related PTK Ce
29	569.5	47.5	849	1	R75706	Eph-related PTK Ce
30	536	44.7	977	1	W19258	Embryonic stem cel
31	535	44.6	951	1	R75704	Eph-related CEK6.
32	466	38.8	522	1	R76466	Non-differentiated
33	466	38.8	522	1	R94653	Receptor type tyro
34	466	38.8	522	1	W06330	Receptor-type tyro
35	466	38.8	522	1	W11303	Receptor-type tyro
36	466	38.8	972	1	R76468	Mature non-differe
37	466	38.8	972	1	W06331	Full-length recept
38	466	38.8	987	1	R85930	Protein tyrosine-k
39	466	38.8	987	1	R89263	Human non-differen
40	466	38.8	987	1	R94652	Receptor type tyro
41	466	38.8	987	1	W06335	Full length recept
42	466	38.8	987	1	W11304	Receptor-type tyro
43	466	38.3	1006	1	W72256	Human receptor typ
44	460	38.3	1006	1	W70525	Human thymus recep
45	460	38.3	1021	1	W70526	Human thymus recep

ALIGNMENTS

RESULT 1
R31466

ID   R31466 standard; Protein; 983 AA.
AC   R31466;
DT   24-MAY-1993 (first entry)
DE   HEK polypeptide.
KW   HEK polypeptide.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FH
FT   peptide   1..20
FT   /note= "Signal peptide"
FT   protein   21..39
FT   /note= "Purified HEK protein #1"
FT   protein   840..860
FT   /note= "Purified HEK protein #2"
FT   modified_site   232..234
FT   /note= "N-link glycosylation"
FT   modified_site   337..339
FT   /note= "N-link glycosylation"
FT   modified_site   391..393
FT   /note= "N-link glycosylation"
FT   modified_site   404..406
FT   /note= "N-link glycosylation"
FT   modified_site   493..495
FT   /note= "N-link glycosylation"
FT   domain   542..565
FT   /note= "Transmembrane domain"
FT   binding_site   628
FT   /note= "ATP binding site"
FT   binding_site   630
FT   /note= "ATP binding site"
FT   binding_site   633
FT   /note= "ATP binding site"
FT   modified_site   779
FT   /note= "Putative autophosphorylation site"
DR   WO9300425-A.
PN   07-JAN-1993
PD   19-JUN-1992; AU0294.
PF   21-JUN-1991; AU-006841.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   WPI; 93-036373/04.
PD   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III'A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence represents human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the

CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence 983 AA;

Query Match   100.0%; Score 1200; DB 1; Length 983;
Best Local Similarity   100.0%;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||||||||||||||||||||||||||||||||||||||
Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271

RESULT 2
ID   R75711 standard; Protein; 983 AA.
AC   R75711;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek4.
OS   Cek4; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis.
OS   Gallus sp.
WO   WO9515375-A.
PN   08-JUN-1995.
PD   08-JUN-1995.
PF   07-SEP-1994;
AC   03-DEC-1993; U10140.
PR   (LJOL-) LA JOLLA CANCER RES FOUND.
PA   Pasquale EB, Sajjadi FG;
PI   WPI; 95-215256/28.
DR   N-PSDB; Q90659.
DR   N-PSDB; Q90659.
PA   (REGE-) REGENERON PHARM INC.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Disclosure; Page 85-89; 129pp; English.
CC   Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and
CC   embryonic tissues and also in the adult brain and retina.
SQ   Sequence 983 AA;

Query Match   94.7%; Score 1136; DB 1; Length 983;
Best Local Similarity   94.1%;
Matches 207; Conservative   7; Mismatches   6; Indels   0; Gaps   0;

Qy   1 WEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPL 60
          ||||||||||||||||||||| ||||||||||||| |||||||||||||||||||||||
Db   51 WEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPL 110

Qy   61 VLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||:||
Db   111 VLGTCKETFNLYYMESDDHLAKFREHQFTKIDTMAADESFTQMDLGDRILKLNTEVREV 170

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db   171 GPVSKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||||||||||||||||||||||||||||:||||||||:
Db   171 GPVSKKGFYLAFQDVGACVALVSVRVYFYKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 230

Db   231 NNSKEEPPRMYCSTEGEWLVPIGKCLCNAGYEERGFACQ 270

RESULT 3
ID   W83148 standard; Protein; 948 AA.
AC   W83148;
DT   11-FEB-1999 (first entry)
DE   Rat receptor tyrosine kinase Ehk-2.
OS   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
OS   binding protein; BDNF; NT-3; diagnosis.
OS   Rattus sp.
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1996; US-469537.
PF   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-735559.
PR   28-OCT-1993; US-144592.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PA   Maisonpierre PC, Masiakowski P, Yancopoulos GD;
DR   N-PSDB; V70208.
WPI; 99-044584/04.
PS   Example; Fig 21; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The present sequence represents rat Ehk-2.
CC   tyrosine kinases. The proteins are orphan receptor
SQ   Sequence 948 AA;

Query Match   76.2%; Score 914.5; DB 1; Length 948;
Best Local Similarity   74.0%;
Matches 159; Conservative 31; Mismatches 24; Indels   1; Gaps   1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
          |:| |||:|||||:||| |||||||||||:|||||::|||||||||||||||||||||||
Db   56 WDAIEMDEHNRPIHIYQVCNVMEPNQNNWLRTNWISRDAAQKIVEMKFTLRDCNSIPW 115

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||:||
Db   116 VLGTCKETFNLYYIESDESHGTKFKFRPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREV 175

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
          ||||||||||||||||||||||||||| |:|| ||||||||||||| ||||||||||||
Db   176 GPIERGFYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSC 235

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 214
          |||||||| :|||||||||||||||||| ||||||
Db   236 VKSSEERDTPRLYCGADGDWNVPLGRCICTTGYEE 270

RESULT 4
ID   R85936 standard; Protein; 986 AA.
AC   R85936;
DT   14-FEB-1996 (first entry)
DE   Protein tyrosine-kinase bpTK7.
KW   Protein tyrosine-kinase; PTK; bpTK7; agonist; cell growth;
KW   differentiation.
OS   Homo sapiens.

FH Key	Location/Qualifiers
FT peptide	1..19
FT	/label= Sig_peptide
FT domain	20..547
FT	/label= Extracellular_domain
FT domain	548..570
FT	/label= Transmembrane_domain
FT domain	571..986
FT	/label= Intracellular_tyrosine-kinase_domain

Db   231 NHSKEEPPKMYCSTEGEWLVPIGKCLCNAGYEERGFACQ 270

PN   WO9527061-A1.
PD   12-OCT-1995;
PF   04-APR-1995; U04228.
PR   04-APR-1994; US-222616.
PA   (GETH) GENENTECH INC.
PA   Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI   Wood WI;
DR   WPI; 95-366160/47.
DR   N-PSDB; T03100.
PT   Agonist antibodies which activate specific protein tyrosine
PT   kinase(s) - also activate chimeric proteins of kinase extracellular
PT   domain and Ig constant domain, useful for studying, and therapeutic
PT   modulation of, cell growth and differentiation
PS   Disclosure; Page 95-99; 125pp; English.
CC   DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC   to screen cDNA libraries to identify novel pTK genes. The bpTKs,
CC   bpTK1, bpTK2, bpTK3, bpTK4, bpTK5 and bpTK7 (R85924-28 and R85935,
CC   respectively) are expressed in human brain tissue and show homology
CC   to known pTKs. A full-length sequence for bpTK7 (R85936) was also
CC   obtd. bpTK7 can be used to design drugs that modulate pTK activity.
SQ   Sequence 986 AA;

Query Match   75.58%; Score 905.5; DB 1; Length 986;
Best Local Similarity   73.3%;   Pred. No. 1.4e-87;
Matches 162; Conservative 30; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:| :|| |||||||||||||:|| |||||||:|: |:||||||||||:||||||:|
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRIDWITREGAQRYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |:||||||||:| ||||:|||||||||||||||||||||||||:||||||:||||:|
Db 114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:|||||||||||||||||||||||||||||||||||||||||| |:||||||||||
Db 174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVRGSC 233

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||| |||:||||||||||| |||||||||||||||
Db 234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 5
R85091
ID   R85091 standard; Protein; 986 AA.
AC   R85091;
DT   16-APR-1996 (first entry)
DR   WPI; 95-37379/48.
DR   N-PSDB; T02948.
DE   New nucleic acid encoding EPH-like receptor protein tyrosine kinase(s) -
DE   and related vectors, host cells, proteins, antibodies etc., used
DE   diagnostically and therapeutically to modulate receptor activation
DE   or prodn.
KW   EPH-like receptor protein tyrosine kinase; HEK8;
KW   human eph-like kinase; therapy; diagnosis; vector; antibody.
OS   Homo sapiens.
PN   WO9528484-A1.
PD   26-OCT-1995;
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PA   Fox GM, Jing S, Welcher AA;
PS   Claim 18; Page 62-65; 133pp; English.
CC   4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC   HEK8 and HEK11 (R85089-92), respectively, were identified following
CC   isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC   cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC   HEK receptors (or their soluble extracellular domains) are produced by

CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   assays), or to identify and purify ligands for HEK receptors, or
CC   therapeutically to modulate the activation of cell-associated
CC   receptors.
SQ   Sequence 986 AA;

Query Match   75.58%; Score 905.5; DB 1; Length 986;
Best Local Similarity   73.3%;   Pred. No. 1.4e-87;
Matches 162; Conservative 30; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:| :|| |||||||||||||:|| |||||||:|: |:||||||||||:||||||:|
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRIDWITREGAQRYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |:||||||||:| ||||:|||||||||||||||||||||||||:||||||:||||:|
Db 114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:|||||||||||||||||||||||||||||||||||||||||| |:||||||||||
Db 174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVRGSC 233

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||| |||:||||||||||| |||||||||||||||
Db 234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 6
W71628
ID   W71628 standard; Protein; 877 AA.
AC   W71628;
DT   25-NOV-1998 (first entry)
DE   Mouse Bsk receptor-like tyrosine kinase.
KW   Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW   neurodegenerative disease; limbic system neuron regeneration;
KW   chromosomal abnormality; degenerative growth; development disorder;
KW   viral infection; bacterial infection; Alzheimer's disease; epilepsy;
KW   schizophrenia; stroke; cerebral ischaemia.
OS   Mus sp.
PN   US5814479-A.
PD   29-SEP-1998;
PF   11-JUN-1996; 673789;
PR   04-JAN-1994; US-177812.
PR   11-JUN-1996; US-673789.
PA   (KROM/) KROMER L F.
PA   (SCHU/) SCHULZ N T.
PA   (WOUD/) WOUDE G F V.
PA   (ZHOU/) ZHOU R.
PI   Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR   WPI; 98-541751/46.
DR   N-PSDB; V58192.
PT   Isolated nucleic acid sequence encoding protein - used in Bsk
PT   nucleic acid probes, used in detecting alterations in level of Bsk
PT   messenger-RNA in biological samples isolated from mammal afflicted
PT   with disease
PS   Claim 1; Fig 2; 7pp; English.
CC   The present sequence represents mouse Bsk, which is a receptor-like
CC   tyrosine kinase. The nucleic acid sequence encoding Bsk can be used in
CC   Bsk nucleic acid probes, which can be used in detecting alterations in
CC   the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC   from a mammal afflicted with a disease, such as neurodegenerative
CC   diseases or disorders and neoplasms. The nucleic acid sequence can also
CC   be delivered into the limbic system of patients with limbic system
CC   neurodegenerative disease, disorder or injury, to promote or enhance
CC   limbic system neuron regeneration or growth. Such neurodegenerative
CC   diseases include, chromosomal abnormalities, degenerative growth and
CC   development disorders, viral infections, bacterial infections, brain
CC   injuries, neoplastic conditions, Alzheimer's disease, epilepsy,
CC   schizophrenia, or stroke and cerebral ischaemia.
SQ   Sequence 877 AA;

```
RESULT 7
W83147
ID   W83147 standard; Protein; 1005 AA.
AC   W83147;
DT   11-FEB-1999  (first entry)
DE   Rat receptor tyrosine kinase Ehk-1.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis.
OS   Rattus sp.
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995; 469537.
PR   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Maslakowski P, Yancopoulos GD;
DR   WPI; 99-044584/04.
DR   N-PSDB; V70207.
PT   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Example; Fig 22; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence represents rat Ehk-1.
SQ   Sequence 1005 AA;

  Query Match     72.2%;  Score 866.5;  DB 1;  Length 877;
  Best Local Similarity   68.3%;  Pred. No. 1.7e-83;
  Matches  151;  Conservative   35;  Mismatches   34;  Indels    1;  Gaps    1;

Qy         1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
             ||||  ||:|| |||:||||||:||:|:|||||| :||| ||:|||||||||||||||:|
Db        85 WEEIGEVDENYAPIHTYQVCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy        61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
             .||||||||:||:|||||| ||:: :|:| |||||||||||::|:|||||:|||||::|:
Db       145 GLGTCKETFNMYYFESDDDENGRSIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy       121 GPVNKKGFYLAFQDVGACVALVSVRVYRVYFKKCPFTVKNLAMFPDVYP-MDSQSLVEVRGSC 179
             ||:.||||||||||||||:|||||||| :||||||: |:::|:||||   |||:||||:|||
Db       205 GPLSKKGFYLAFQDVGACIALVSVRVYYKKCPSVVRHLAIFPDITGADSSQLLEVSGSC 264

Qy       180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
             ||:|  |||:|:||:|||||||||||| |:|||||: | ||
Db       265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305


RESULT 8
R97853
ID   R97853 standard; Protein; 928 AA.
AC   R97853;
DT   05-JAN-1997  (first entry)
DE   Rat REK7 eph-related tyrosine kinase receptor.
KW   REK7; eph-related tyrosine kinase receptor; AL-1; ligand;
KW   neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   peptide         1..57
FT                   /label= Sig_peptide
FT   protein         58..928
FT                   /label= Mat_protein
FT   domain          58..462
FT                   /label= Extracellular_domain
PN   WO9613518-A1.
PD   09-MAY-1996.
PF   26-OCT-1995; US14016.
PR   27-OCT-1994; US-330128.
PR   07-JUN-1995; US-486449.
PA   (GETH ) GENENTECH INC.
PI   Caras IW, Winslow JW;
DR   WPI; 96-239448/24.
DR   N-PSDB; T18893.
PT   AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
PT   treatment and diagnosis of neuronal disorders and
PT   angiogenesis-related conditions.
PS   Example 1; Page 50-53; 75pp; English.
CC   Rat REK7 (R97853) is an eph-related tyrosine kinase receptor, for
CC   which AL-1 (see also W97854) is a ligand. Its amino acid sequence
CC   was deduced from a cDNA clone (T18893) isolated from a hippocampal
CC   cDNA library. An REK7-IgG fusion was used to screen cultured cell
CC   lines for surface expression of REK7-binding activity. Primers
CC   (see also T44382-83) based on peptide sequences (R97856-59) of
CC   isolated ligands were used to amplify human breast carcinoma BT20
CC   cell cDNA, and an amplified fragment was used to screen a human
CC   foetal brain cDNA library, leading to the isolation of AL-1 cDNA
CC   (T18897).
SQ   Sequence 928 AA;

  Query Match     72.1%;  Score 865.5;  DB 1;  Length 928;
  Best Local Similarity   68.3%;  Pred. No. 2.3e-83;
  Matches  151;  Conservative   35;  Mismatches   34;  Indels    1;  Gaps    1;

Qy         1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
             ||||  ||:|| |:||||||:||:|:||||||:: |||| ||:|||||||||||||||:|
Db        85 WEEIGEVDENYPPIHTYQVCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy        61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
             .||||||||:||:||||| :|:|: :|:| |||||||||||::|:|||||:|||||::|:
Db       145 GLGTCKETFNMYYFESDDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy       121 GPVNKKGFYLAFQDVGACVALVSVRVYRVYFKKCPFTVKNLAMFPDVYP-MDSQSLVEVRGSC 179
             ||:.||||||||||||||:|||||||| :||||||: |:::|:||||   |||:||||:|||
Db       205 GPLSKKGFYLAFQDVGACIALVSVRVYYKKCPSVVRHLAIFPDITGADSSQLLEVSGSC 264

Qy       180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
             ||:|  |||:|:||:|||||||||||| |:|||||: | ||
Db       265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305


RESULT 9
R85090
ID   R85090 standard; Protein; 991 AA.
AC   R85090;
DT   16-APR-1996  (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK7.
KW   EPH-like receptor protein tyrosine kinase; HEK7;
KW   human eph-like receptor protein tyrosine kinase; therapy; diagnosis; vector; antibody.
OS   Homo sapiens.

  Query Match     72.8%;  Score 866.5;  DB 1;  Length 877;
  Best Local Similarity   68.3%;  Pred. No. 2.1e-83;
  Matches  151;  Conservative   35;  Mismatches   34;  Indels    1;  Gaps    1;

Qy         1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
             ||||  ||:|| |||:||||||:||:|:|||||| :||| ||:|||||||||||||||:|
Db        85 WEEIGEVDENYAPIHTYQVCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy        61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
             .||||||||:||:||||| :|:|: :|:| |||||||||||::|:|||||:|||||::|:
Db       145 GLGTCKETFNMYYFESDDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy       121 GPVNKKGFYLAFQDVGACVALVSVRVYRVYFKKCPFTVKNLAMFPDVYP-MDSQSLVEVRGSC 179
             ||:.||||||||||||||:|||||||| :||||||: |:::|:||||   |||:||||:|||
Db       205 GPLSKKGFYLAFQDVGACIALVSVRVYYKKCPSVVRHLAIFPDITGADSSQLLEVSGSC 264

Qy       180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
             ||:|  |||:|:||:|||||||||||| |:|||||: | ||
Db       265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305
```

PN WO952844-A1.
PD 26-OCT-1995;
PF 14-APR-1995; U04681.
PR 15-APR-1994; US-229509.
PA (AMGE-) AMGEN INC.
PI Fox GM, Jing S, Welcher AA;
DR N-PSDB; T02947.
DR WPI; 95-37799/48.
PT New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT and related vectors, host cells, proteins, antibodies etc., used
PT diagnostically and therapeutically to modulate receptor activation
PT or prodn.
PS Claim 18; Page 54-57; 133pp; English.
CC 4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC HEK8 and HEK11 (R85089-92), respectively, were identified following
CC isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC HEK receptors (or their soluble extracellular domain) are produced by
CC expression of encoding sequences in procaryotic or eucaryotic host
CC cells, and are used to produce antibodies (utilised in diagnostic
CC assays), or to identify and purify ligands for HEK receptors, or
CC therapeutically to modulate the activation of cell-associated
CC receptors. Soluble HEK7 receptor may primarily affect
CC proliferation and/or differentiation of brain cells.
CC pancreatic cells.
SQ Sequence 991 AA;

Qy Query Match   71.7%; Score 860.5; DB 1; Length 991;
    Best Local Similarity 67.9%; Pred. No. 8.7e-83;
    Matches 150; Conservative 35; Mismatches 35; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||| ||:||||||||||||||||| |: |||||||:| ||||||||||||||||||
Db   59 WEEIGEVDENYAPIHTYQVCKVWMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 118

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         :|||||||:|| |||||::|||:|||:||||||||||||||||||||:||:|||||::|:|
Db 119 GLGTCKETFNMYYFESDDQNGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 178

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:|||||||||||||||:||||||||:|||||||:|||||||:||| |||||||||||||
Db 179 GPLSKKGFYLAFQDVGACIALVSVRVYIKKCPSVVRIHLAVFPDIITGADSSQLEVSGSC 238

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||| ||::|| :|||||||||||||||||||||||||||||
Db 239 VNHSVIDEPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 279

RESULT 10
ID W03422 standard; Protein; 610 AA.
AC W03422;
DT 11-NOV-1996 (first entry)
DE Mouse developmental kinase 1 MDK1 T1.
KW Mouse developmental kinase 1; MDK1 T1; receptor tyrosine kinase;
KW RTK; signal transduction; probe; diagnosis; gene therapy;
KW neurodegeneration; neuroproliferation; cancer.
OS Mus sp.
FH Key   Location/Qualifiers
FT peptide   1..28
FT   /label= Sig_peptide
FT modified_site   64..66
FT   /label= N-glycosylation_site
FT modified_site   343..345
FT   /label= N-glycosylation_site
FT modified_site   410..412
FT   /label= N-glycosylation_site
FT domain   555..579
FT   /label= Transmembrane_domain
FT misc_difference 600..610

FT   /note= "product of alternative splicing"
PN WO9621013-A1.
PD 11-JUL-1996;
PF 03-JAN-1995; U00419.
PR 03-JAN-1995; US-368776.
PA (PLAC-) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (SUGE-) SUGEN INC.
PI Ciossek T, Millauer B, Ullrich A;
DR N-PSDB; T32962.
DR WPI; 96-333988/33.
PT New mouse development kinase 1 gene - used for developing prods. for
PT diagnosis and treatment of abnormalities in signal transduction
PT pathways
PS Example 1; Page 109-111; 128pp; English.
CC cDNA cloning using adult mouse brains and Northern blotting
CC identified 2 truncated versions, MDK1 T1 (W03422) and MDK1 T2
CC (W03423), of the novel mouse developmental kinase 1 (MDK1) (see also
CC W03421), a new member of the eck/eph family of receptor tyrosine
CC kinases. Their amino acid sequences were deduced from cDNA clones
CC (T32961 and T32962) obtd. from adult mouse brains. MDK T1 and T2
CC each possess the entire ectodomain, the transmembrane domain and
CC part of the juxtamembrane region of MDK1, but lack the catalytic
CC tyrosine kinase domain. They can be used to screen for potential
CC agents useful for treatment of diseases characterised by abnormal
CC signal transduction.
SQ Sequence 610 AA;

Qy Query Match   71.5%; Score 857.5; DB 1; Length 610;
    Best Local Similarity 70.0%; Pred. No. 9.2e-83;
    Matches 156; Conservative 36; Mismatches 28; Indels 3; Gaps 2;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||| |:|:||||||||||||: |||||||||||||: ||||:||||||||||||||:|
Db   55 WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||||| :|:|||:|| |||:|||||||||||||:||||:|:||||||:||:|
Db 115 VLGTCKETFNLYYETDYDTGRNIRENLYVKIDTIAADESFTQSDLGERMKLNTEVREI 174

Qy 121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||:|||||||||||||||:||||:||||:||| :||||||||:|| ||||||||||||
Db 175 GPLSKKGFYLAFQDVGACIALVSVKVYIKKCWIIVENLAVFPDIVTGSEFSLVEVRGTC 234

Qy 180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |:::|: ||:|||:||||||||||||||||||||||| :||||
Db 235 VSSAEEEAENSPRMHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277

RESULT 11
ID W03423 standard; Protein; 626 AA.
AC W03423;
DT 11-NOV-1996 (first entry)
DE Mouse developmental kinase 1 MDK1 T2.
KW Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW RTK; signal transduction; probe; diagnosis; gene therapy;
KW neurodegeneration; neuroproliferation; cancer.
OS Mus sp.
FH Key   Location/Qualifiers
FT peptide   1..28
FT   /label= Sig_peptide
FT modified_site   64..66
FT   /label= N-glycosylation_site
FT modified_site   343..345
FT   /label= N-glycosylation_site
FT modified_site   410..412
FT   /label= N-glycosylation_site
FT domain   555..579
FT   /label= Transmembrane_domain
FT misc_difference 600..626
FT   /note= "product of alternative splicing"
PN WO9621013-A1.

PD 11-JUL-1996.
PF 03-JAN-1996; U00419.
PR 03-JAN-1995; US-368776.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN
PA (SUGE-) SUGEN INC.
PI Clossek T, Millauer B, Ullrich A;
DR WPI; 96-333988/33.
DR N-PSDB; T32962.
PN Example 1; Page 113-115; 128pp; English.
PT New mouse development kinase 1 gene - used for developing prods. for
PT diagnosis and treatment of abnormalities in signal transduction
PT pathways
CC Example 1; Page 113-115; 128pp; English.
CC cDNA cloning using adult mouse brains and Northern blotting
CC identified 2 truncated versions, MDK1 T1 (W03422) and MDK1 T2
CC (W03423), of the novel mouse developmental kinase 1 (MDK1) (see also
CC W03421), a new member of the eck/eph family of receptor tyrosine
CC kinases. Their amino acid sequences were deduced from cDNA clones
CC (T32961 and T32962) obtd. from adult mouse brains. MDK T1 and T2
CC each possess the entire ectodomain, the transmembrane domain and
CC part of the juxtamembrane region of MDK1, but lack the catalytic
CC tyrosine kinase domain. They can be used to screen for potential
CC agents useful for treatment of diseases characterised by abnormal
CC signal transduction.
SQ Sequence   626 AA;

Query Match   71.5%; Score 857.5; DB 1; Length 626;
Best Local Similarity   70.0%; Pred. No. 9.5e-83;
Matches 156; Conservative 36; Mismatches 28; Indels   3; Gaps   2;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||| ||::||||||||||||||:||||||||||||||||
Db   55 WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNIWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||| |||::||||||||||||||||||||||||| |||:|||||||||||:|
Db 115 VLGTCKETFNLYYETDYDTGRNIRENLVYKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         |||||||:|||||||||||||:||:||||||||||||||||||||||||| :|||:||||||:|
Db 175 GPLSKKGFYLAFQDVGACIALVSVAVYYKRCWTIVENLAVFPDTVTGSEFSSLVEVRGTC 234

Qy 180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         :::::| ||||||||||||||||||||||||||||||||:||
Db 235 VSSAEEEAENSPRHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277

RESULT 12
ID W03421 standard; Protein; 998 AA.
AC W03421;
DT 11-NOV-1996 (first entry)
DE Mouse developmental kinase 1.
KW Mouse developmental kinase 1; MDK1; receptor tyrosine kinase; RTK;
KW signal transduction; probe; diagnosis; therapy;
KW neurodegeneration; neuroproliferation; cancer.
OS Mus sp.
FH Key   Location/Qualifiers
FT peptide   1..28
FT   /label= sig_peptide
FT modified_site   64..66
FT   /label= N-glycosylation_site
FT modified_site   343..345
FT   /label= N-glycosylation_site
FT modified_site   410..412
FT   /label= N-glycosylation_site
FT modified_site   555..579
FT   /label= N-glycosylation_site
FT domain   555..579
FT   /label= Transmembrane_domain
FT   /label= domain
FT WO962I013-A1.
FF 11-JUL-1996; U00419.
PD 03-JAN-1996; U00419.
PR 03-JAN-1995; US-368776.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (SUGE-) SUGEN INC.
PI Clossek T, Millauer B, Ullrich A;
PA WPI; 96-333988/33.
DR N-PSDB; T32960.
PN New mouse development kinase 1 gene - used for developing prods. for
PT diagnosis and treatment of abnormalities in signal transduction
PT pathways
PS Example 1; Page 105-108; 128pp; English.
CC Mouse developmental kinase 1 (MDK1) (W03421) is a new member of the
CC eck/eph family of receptor tyrosine kinases (RTKs). Its amino
CC acid sequence was deduced from a cDNA clone (T32960) isolated
CC from mouse embryo and adult brain libraries. The distinct
CC patterns of MDK1 expression during mouse development suggest an
CC important role for MDK1 in the formation of neuronal structures.
CC MDK1 may be obtd. by expression in host cells. It can be used
CC in methods for the diagnosis of diseases characterised by
CC abnormality in a signal transduction pathway, such as
CC neuroproliferative or neurodegenerative disorders or cancer,
CC to screen for (ant)agonists, and to raise antibodies.
SQ Sequence   998 AA;

Query Match   71.5%; Score 857.5; DB 1; Length 998;
Best Local Similarity   70.0%; Pred. No. 1.8e-82;
Matches 156; Conservative 36; Mismatches 28; Indels   3; Gaps   2;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||| ||::||||||||||||||:||||||||||||||||
Db   55 WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNIWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||| |||::||||||||||||||||||||||||| |||:|||||||||||:|
Db 115 VLGTCKETFNLYYETDYDTGRNIRENLVYKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         |||||||:|||||||||||||:||:||||||||||||||||||||||||| :|||:||||||:|
Db 175 GPLSKKGFYLAFQDVGACIALVSVAVYYKRCWTIVENLAVFPDTVTGSEFSSLVEVRGTC 234

Qy 180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         :::::| ||||||||||||||||||||||||||||||||:||
Db 235 VSSAEEEAENSPRHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277

RESULT 13
ID R85092
AC R85092;
DT 16-APR-1996 (first entry)
DE EPH-like receptor protein tyrosine kinase HEK11.
KW EPH-like receptor protein tyrosine kinase; PTK; HEK11;
KW human eph-like kinase; therapy; diagnosis; antibody; vector.
OS Homo sapiens.
PN WO952B484-A1.
PD 26-OCT-1995.
PF 14-APR-1995; U04681.
PR 15-APR-1994; US-229509.
PA (AMGE-) AMGEN INC.
PI Fox GM, Jing S, Welcher AA;
DR WPI; 95-373799/48.
DR N-PSDB; T02949.
PT New nucleic acid encoding EPH-like receptor protein tyrosine kinase(s)
PT and related vectors, host cells, proteins, antibodies etc., used
PT diagnostically and therapeutically to modulate receptor activation
PT or prodn.
PS Claim 18; Page 71-75; 133pp; English.
CC 4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC HEK8 and HEK11 (R85089-92), respectively, were identified following
CC isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC HEK receptors (or their soluble extracellular domains) are produced by
CC expression of encoding sequences in procaryotic or eucaryotic host

Query Match   71.5%; Score 857.5; DB 1; Length 998;
Best Local Similarity   70.0%; Pred. No. 1.8e-82;
Matches 156; Conservative 36; Mismatches 28; Indels   3; Gaps   2;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||| ||::||||||||||||||:||||||||||||||||
Db   55 WEEISGLDENYTPIRTYQVCQVMEPNQNNWLRTNIWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||| |||::||||||||||||||||||||||||| |||:|||||||||||:|
Db 115 VLGTCKETFNLYYETDYDTGRNIRENLVYKIDTIAADESFTQGDLGERKMKLNTEVREI 174

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         |||||||:|||||||||||||:||:||||||||||||||||||||||||| :|||:||||||:|
Db 175 GPLSKKGFYLAFQDVGACIALVSVAVYYKRCWTIVENLAVFPDTVTGSEFSSLVEVRGTC 234

Qy 180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         :::::| ||||||||||||||||||||||||||||||||:||
Db 235 VSSAEEEAENSPRHCSAEGEWLVPIGKCICKAGYQOKGDTCE 277

```
CC  cells, and are used to produce antibodies (utilised in diagnostic
CC  assays), or to identify and purify ligands for HEK receptors, or
CC  therapeutically to modulate the activation of cell-associated
CC  receptors.
SQ  Sequence  998 AA;

Query Match      71.4%;  Score 856.5;  DB 1;  Length 998;
Best Local Similarity  69.5%;  Pred. No. 2.3e-82;
Matches 155;  Conservative 37;  Mismatches 28;  Indels  3;  Gaps  2;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||||:||||| ||:||||||||||||||||||: ||:|:|||||||||||||:|||||
Db   55 WEEISGLDENYTPIRIYQVCQVMEPNQNNWLRTNWISKGNAQRIFVELKFTLRDCNSLPG 114

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||:|:|| |||:|||||||||||||||||||:|:|||||||||||:|:|
Db  115 VLGTCKETFNLYYEDFDYDTGRNIRENLYVKIDTIAADESFTQGDLGGERKMKINTEVREI 174

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         |:||||||||||||||||||||||||||||||:||||||||||||||  :||||:|||:|
Db  175 GPLSKKGFYLAFQDVGACIALVSVKVYKYKKCWSIIENLAIFPDTVTGSEFSSLVEVRGTC 234

Qy  180 VNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         |:::|| |||||||||||||||||||:||||||:|:::||:|:
Db  235 VSSAEEAENAPRMHCSAEGEWLVPIGKCICKAGYQQKGDTCE 277
```

```
RESULT  14

ID  R75712  standard;  Protein;  995 AA.
AC  R75712;
DT  11-NOV-1995  (first entry)
DE  Eph-related PTK Cek5.
KW  Cek5; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW  prognosis.
OS  Gallus sp.
PN  WO9515375-A.
PD  08-JUN-1995.
PF  07-SEP-1994; U10140.
PR  03-DEC-1993; US-162809.
PA  (LJOL-) LA JOLLA CANCER RES FOUND.
PI  Pasquale EB, Sajjadi FG;
PT  Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT  cancer.
DR  N-PSDB; Q90660.
DR  WPI; 95-215256/28.
PS  Disclosure; Page 92-96; 129pp; English.
CC  Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
CC  (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC  Q90661-62) from chicken embryo and embryonic brain libraries.
SQ  Sequence  995 AA;

Query Match      59.4%;  Score 713;  DB 1;  Length 995;
Best Local Similarity  56.1%;  Pred. No. 4e-67;
Matches 128;  Conservative  48;  Mismatches  44;  Indels   8;  Gaps   4;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:||:|| ||:||||||: |||||||||||||:|: :|||||:|||||:|||||||||
Db   51 WEEVSGYDENMNTIRIYQVCNVFESSQNNWLRTKYIRRGAHRIHVEMKFSVRDCSSIPN 110

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTE 116
         |:||||||||||| ||| |     |||||||||||||||||||||||:|:||||||||
Db  111 VPGSCKETFNLYYESDFDSATKTFPNWMENPWMKVDTIAADESFSQVDLGGRVMKINTE 170

Qy  117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         :| ||||||||||||||||||||||||||:|||||||:||||||||||||  |||||:||
Db  171 VRSFGPVSKNGFYLAFQDYGGCMSLIAVRVFYRKCPRVIQNGAVFQETLSGAESTSLVA 230

Qy  176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE--ERGFMCQ 220
         :|||:|::|:| |:|||||||||||||||||:|||:||||  :||:|:
Db  231 RGTCISNAEEVDVPIKLYCNDGEWLVPIGRCMCRPGYESEVENGTVCR 278
```

```
RESULT  15

ID  R75709  standard;  Protein;  1011 AA.
AC  R75709;
DT  11-NOV-1995  (first entry)
DE  Eph-related PTK Cek5+.
KW  Cek5+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW  prognosis.
OS  Gallus sp.
PN  WO9515375-A.
PD  08-JUN-1995.
PF  07-SEP-1994; U10140.
PR  03-DEC-1993; US-162809.
PA  (LJOL-) LA JOLLA CANCER RES FOUND.
PI  Pasquale EB, Sajjadi FG;
PT  Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT  cancer.
DR  WPI; 95-215256/28.
DR  N-PSDB; Q90657.
PS  Claim 11; Page 71-75; 129pp; English.
CC  A cDNA clone encoding a novel variant of EPH-related PTK Cek5,
CC  Cek5+ (Q90657), was isolated from a chick embryo library in
CC  lambda gt11. Cek5+ protein (R75709) contains a 16-amino acid
CC  insertion in the juxtamembrane domain, and be a result of
CC  alternative splicing. Cek5+ is exclusively expressed in the CNS.
SQ  Sequence  1011 AA;

Query Match      59.4%;  Score 713;  DB 1;  Length 1011;
Best Local Similarity  56.1%;  Pred. No. 4.1e-67;
Matches 128;  Conservative  48;  Mismatches  44;  Indels   8;  Gaps   4;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:||:|| ||:||||||: |||||||||||||:|: :|||||:|||||:|||||||||
Db   51 WEEVSGYDENMNTIRIYQVCNVFESSQNNWLRTKYIRRGAHRIHVEMKFSVRDCSSIPN 110

Qy   61 VLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTE 116
         |:||||||||||| ||| |     |||||||||||||||||||||||:|:||||||||
Db  111 VPGSCKETFNLYYESDFDSATKTFPNWMENPWMKVDTIAADESFSQVDLGGRVMKINTE 170

Qy  117 IREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEV 175
         :| ||||||||||||||||||||||||||:|||||||:||||||||||||  |||||:||
Db  171 VRSFGPVSKNGFYLAFQDYGGCMSLIAVRVFYRKCPRVIQNGAVFQETLSGAESTSLVAA 230

Qy  176 RGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE--ERGFMCQ 220
         :|||:|::|:| |:|||||||||||||||||:|||:||||  :||:|:
Db  231 RGTCISNAEEVDVPIKLYCNDGEWLVPIGRCMCRPGYESEVENGTVCR 278
```

```
Search completed: May  9, 2000, 21:52:01
Job time: 3753 sec
```

This Page Blank (uspto)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 22:31:07 ; Search time 71.69 Seconds
                                              (without alignments)
                                              21.277 Million cell updates/sec

Title: US-09-104-340-2
Perfect score: 119
Sequence: 1 VNLLDSKTIQGELGWISYPSHG 22
Scoring table: BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched: 225878 seqs, 69334122 residues
Total number of hits satisfying chosen parameters: 225878
Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 listing first 45 summaries

Database :   SPTREMBL_12:*
                     1: sp_archea:*
                     2: sp_bacteria:*
                     3: sp_fungi:*
                     4: sp_human:*
                     5: sp_invertebrate:*
                     6: sp_mammal:*
                     7: sp_mhc:*
                     8: sp_organelle:*
                     9: sp_phage:*
                     10: sp_plant:*
                     11: sp_rodent:*
                     12: sp_virus:*
                     13: sp_vertebrate:*
                     14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	79	66.4	985	13	Q91694	Q91694 xenopus lae
2	79	66.4	986	13	Q91845	Q91845 xenopus lae
3	71	59.7	981	13	096993	096993 heterodera
4	71	59.7	981	13	O13146	O13146 brachydanio
5	61	51.3	993	13	042422	042422 gallus gall
6	58	48.7	985	13	Q91571	Q91571 xenopus lae
7	57	47.9	974	13	Q91735	Q91735 xenopus lae
8	56	47.1	1055	4	043477	043477 homo sapien
9	55	46.2	984	13	043569	043569 homo sapien
10	55	46.2	984	4	095142	095142 homo sapien
11	53	44.5	952	13	Q07494	Q07494 gallus gall
12	52	43.7	1006	11	015197	015197 homo sapien
13	51	42.9	1014	4	008644	008644 mus musculu
14	51	42.9	366	2	Q50031	Q50031 mycobacteri
15	49	41.2	368	2	053825	053825 mycobacteri
16	48	40.3	1763	5	Q17901	Q17901 caenorhabdi
17	48	40.3	1769	5	Q9XTJ1	Q9XTJ1 caenorhabdi
18	46	38.7	134	9	038422	038422 bacteriopha
19	46	38.7	217	5	061063	061063 physarum po
20	45	37.8	809	10	Q92PC6	Q92PC6 craterostig
21	45	37.8	90	2	033642	033642 salmonella
22	45	37.8	464	5	Q21794	Q21794 caenorhabdi
23	45	37.8	530	12	083884	083884 norwalk vir
24	45	37.8	1151	5	096993	096993 heterodera
25	44	37.0	89	2	P71504	P71504 methylobact
26	44	37.0	234	2	Q92CV2	Q92CV2 rickettsia
27	44	37.0	540	12	Q9YS14	Q9YS14 norwalk vir
28	44	37.0	794	12	Q66530	Q66530 european br
29	43	36.1	2334	12	096725	096725 european br
30	43	36.1	338	2	Q59082	Q59082 azospirillu
31	43	36.1	360	5	Q18228	Q18228 caenorhabdi
32	43	36.1	489	3	014237	014237 schizosacch
33	43	36.1	435	10	023393	023393 arabidopsis
34	43	36.1	652	5	060993	060993 crithidia f
35	43	36.1	1364	4	Q9YSP7	Q9YSP7 homo sapien
36	43	36.1	1368	4	Q9Y667	Q9Y667 homo sapien
37	42.5	35.7	85	1	028496	028496 archaeoglob
38	42.5	35.7	537	2	005459	005459 sciara ocel
39	42.5	35.7	1252	5	077403	077403 mycobacteri
40	42	35.3	112	12	Q9YR39	Q9YR39 aleutian ml
41	42	35.3	112	12	Q9YR38	Q9YR38 aleutian ml
42	42	35.3	217	8	099544	099544 paramesopod
43	42	35.3	358	2	034753	034753 bacillus su
44	44	35.3	627	5	015832	015832 leishmania
45	41.5	34.9	284	5	017213	017213 caenorhabdi

ALIGNMENTS

RESULT 1

ID   Q91694   PRELIMINARY;   PRT;   985 AA.
AC   Q91694;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR PAG PRECURSOR (EC 2.7.1.112)
DE   (PAGLIACCIO).
GN   PAG.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NEURAL CREST;
RX   MEDLINE; 95001564.
RA   WINNING R.S., SARGENT T.D.;
RT   "Pagliaccio, a member of the Eph family of receptor tyrosine kinase
RT   genes, has localized expression in a subset of neural crest and neural
RT   tissues in Xenopus laevis embryos.";
RL   Mech. Dev. 46:219-229(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY
CC   SIMILARITY). MAY PLAY A ROLE IN THE DIFFERENTIATION OF CRANIAL
CC   NEURAL CREST AND OTHER TISSUES.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: LOCALIZED EXPRESSION IN A SUBSET OF NEURAL
CC   CREST AND NEURAL TISSUES IN EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: PRESENT TRANSIENTLY IN VISCERAL ARCH 3. ALSO
CC   EXPRESSED IN THE FOREBRAIN, RHOMBOMERES R3 AND R5 OF THE HINDBRAIN
CC   AND IN THE PRONEPHROS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; L26099; AAA64464.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.

Left column (RESULT 2 record — continued)

```
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
DR   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL          1     20       POTENTIAL.
FT   CHAIN          21    985       TYROSINE-PROTEIN KINASE RECEPTOR PAG.
FT   DOMAIN         21    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      548    569       POTENTIAL.
FT   DOMAIN        570    985       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         66    115       IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN        191    325       CYS-RICH.
FT   DOMAIN        326    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN        436    533       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN        620    881       PROTEIN KINASE.
FT   NP_BIND       626    634       ATP (BY SIMILARITY).
FT   DISULFID       73    108       BY SIMILARITY.
FT   BINDING       652    652       ATP (BY SIMILARITY).
FT   ACT_SITE      745    745       BY SIMILARITY.
FT   MOD_RES       778    778       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD      340    340       POTENTIAL.
FT   CARBOHYD      407    407       POTENTIAL.
SQ   SEQUENCE      985 AA;  109729 MW;  026BA8A5 CRC32;

Qy    1  VNLLDSKTIQGELGWISYPSHG  22
         ||||:|:||:||||||||:||||
Db   31  VTLLDSRSVQGELGWIASPLEG  52

Query Match           66.4%;  Score 79;  DB 13;  Length 985;
Best Local Similarity  63.6%;  Pred. No. 0.00037;
Matches   14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

RESULT   2
ID   Q91845                     PRELIMINARY;    PRT;    986 AA.
AC   Q91845;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR SEK-1 PRECURSOR (EC 2.7.1.112).
GN   SEK-1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96125143.
RA   XU Q., ALLDUS G., HOLDER N., WILKINSON D.G.;
RT   "Expression of a truncated Sek-1 receptor tyrosine kinase disrupts the
RT   segmental restriction of gene expression in the Xenopus and zebrafish
RT   hindbrain.";
RL   Development 121:4005-4016(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY
CC       SIMILARITY). REQUIRED FOR INTERACTIONS THAT REGULATE THE SEGMENTAL
CC       IDENTITY OR MOVEMENT OF CELLS.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSION OCCURS IN R3, R5 AND TRANSIENTLY
CC       AT LOWER LEVELS IN R2.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; X91191; CAA62601.1; -.
```

Right column (RESULT 3 record)

```
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
DR   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL          1     15       POTENTIAL.
FT   CHAIN          16    986       TYROSINE-PROTEIN KINASE RECEPTOR SEK-1.
FT   DOMAIN         16    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      548    569       POTENTIAL.
FT   DOMAIN        570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         66    115       IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN        191    325       CYS-RICH.
FT   DOMAIN        326    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN        436    533       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN        620    881       PROTEIN KINASE.
FT   NP_BIND       626    634       ATP (BY SIMILARITY).
FT   DISULFID       73    108       BY SIMILARITY.
FT   BINDING       652    652       ATP (BY SIMILARITY).
FT   ACT_SITE      745    745       BY SIMILARITY.
FT   MOD_RES       778    778       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD      340    340       POTENTIAL.
FT   CARBOHYD      407    407       POTENTIAL.
SQ   SEQUENCE      986 AA;  109840 MW;  028DBBDA CRC32;

Qy    1  VNLLDSKTIQGELGWISYPSHG  22
         ||||:|:||:||||||||:||||
Db   31  VTLLDSRSVQGELGWIASPLEG  52

Query Match           66.4%;  Score 79;  DB 13;  Length 986;
Best Local Similarity  63.6%;  Pred. No. 0.00037;
Matches   14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

RESULT   3
ID   O13146                     PRELIMINARY;    PRT;    981 AA.
AC   O13146;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE KINASE 1 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN KINASE
DE   RECEPTOR ZEK1).
GN   ZEK1.
OS   Brachydanio rerio (zebrafish) (zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC   Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC   Cyprinoidea; Cyprinidae; Rasborinae; Danio.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97329578.
RA   BOVENKAMP D.E., GREER P.;
RT   "Novel Eph-family receptor tyrosine kinase is widely expressed in the
RT   developing zebrafish nervous system.";
RL   Dev. Dyn. 209:166-181(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC       ROLE IN EARLY PATTERN FORMATION WITHIN THE DEVELOPING NERVOUS
CC       SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN THE DEVELOPING ZEBRAFISH
CC       NERVOUS SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
```

```
RESULT 4
ID   O42422      PRELIMINARY;  PRT;  993 AA.
AC   O42422;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE RECEPTOR (EC 2.7.1.112) (TYROSINE-
DE   PROTEIN KINASE RECEPTOR CEPHA7).
GN   CEPHA7.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 98092111.
RA   ARAUJO M., NIETO M.A.;
RT   "The expression of chick Epha7 during segmentation of the central and
RT   peripheral nervous system.";
RL   Mech. Dev. 68:173-177(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U89295; AAC60220.1; -.
DR   HSSP; P00523; 2PTK.
DR   ZFIN; ZDB-GENE-990415-58; zek1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Receptor; Transferase; Tyrosine-protein kinase; ATP-binding;
KW   Phosphorylation; Transmembrane; Glycoprotein; Signal;
KW   Immunoglobulin domain.
FT   SIGNAL       1     20       BY SIMILARITY.
FT   CHAIN       21    981       EPH-LIKE KINASE 1.
FT   DOMAIN      21    545       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   546    566       POTENTIAL.
FT   DOMAIN     567    981       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      64    113       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN     192    325       CYS-RICH.
FT   DOMAIN     326    438       FIBRONECTIN TYPE-III.
FT   DOMAIN     439    532       FIBRONECTIN TYPE-III.
FT   DOMAIN     626    887       PROTEIN KINASE.
FT   NP_BIND    632    640       ATP (BY SIMILARITY).
FT   DISULFID    71    106       BY SIMILARITY.
FT   BINDING    658    658       ATP (BY SIMILARITY).
FT   ACT_SITE   751    751       BY SIMILARITY.
FT   MOD_RES    784    784       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD   340    340       POTENTIAL.
FT   CARBOHYD   410    410       POTENTIAL.
FT   CARBOHYD   435    435       POTENTIAL.
FT   CARBOHYD   485    485       POTENTIAL.
FT   VARIANT    141    141       POTENTIAL.
FT   VARIANT    141    141       S -> N.
SQ   SEQUENCE   981 AA;  109654 MW;  10D38182 CRC32;

Qy         1  VNLLDSKTIQGELGWISYPSHG   22
              ||||   :||||  :||||   |
Db        30  VTLLDSMSAPGDLGWEAYPSEG   51

Query Match  59.1%;  Best Local Similarity  59.7%;
Matches  13;  Conservative  3;  Mismatches  6;  Indels  0;  Gaps  0;
Score  71;  Pred. No.  0.0067;  DB  13;  Length  981;
```

```
(database entry)
CC   -!- TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WITHIN THE NERVOUS SYSTEM, EXPRESSION IS
CC       RESTRICTED TO PROSOMERES 1 AND 2 IN THE DIENCEPHALON AND ALL THE
CC       RHOMBOMERES IN THE HINDBRAIN DURING SEGMENTATION STAGES. LATER ON,
CC       A SUPERIMPOSED PATTERN APPEARS THAT CORRELATES WITH THE FORMATION
CC       OF SEVERAL AXONAL TRACTS. IN THE SOMITIC MESODERM, THE EXPRESSION
CC       CORRELATES WITH SEGMENTATION AND THE GUIDANCE OF BOTH NEURAL CREST
CC       AND MOTOR AXONS THROUGH THE SCLEROTOMES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; Y14271; CAA74643.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Receptor; Transferase; Tyrosine-protein kinase; ATP-binding;
KW   Phosphorylation; Transmembrane; Glycoprotein; Signal;
KW   Immunoglobulin domain.
FT   SIGNAL       1     30       POTENTIAL.
FT   CHAIN       31    993       EPH-LIKE RECEPTOR TYROSINE KINASE.
FT   DOMAIN      31    551       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   552    572       POTENTIAL.
FT   DOMAIN     573    993       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      67    116       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN     192    328       CYS-RICH.
FT   DOMAIN     329    438       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN     439    536       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN     628    889       PROTEIN KINASE.
FT   NP_BIND    634    642       ATP (BY SIMILARITY).
FT   DISULFID    74    109       BY SIMILARITY.
FT   BINDING    660    660       ATP (BY SIMILARITY).
FT   ACT_SITE   753    753       BY SIMILARITY.
FT   MOD_RES    786    786       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD   343    343       POTENTIAL.
FT   CARBOHYD   410    410       POTENTIAL.
SQ   SEQUENCE   993 AA;  111366 MW;  6B6BFD17 CRC32;

Qy         1  VNLLDSKTIQGELGWISYPSHG    22
              ||||  :||||| ||| :| :|
Db        33  VILLDSKAQQCTELEWISSPPNG   54

Query Match  63.6%;  Best Local Similarity  51.3%;
Matches  14;  Conservative  1;  Mismatches  7;  Indels  0;  Gaps  0;
Score  61;  Pred. No.  0.26;  DB  13;  Length  993;
```

```
RESULT 5
ID   Q91571      PRELIMINARY;  PRT;  985 AA.
AC   Q91571;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR XEK PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR XEK).
GN   XEK.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95215070.
```

RESULT 6

```
Qy    3  LLDSKTIQGELGWISYPSHG  22
         |:|::| |||| :| |||
Db   24  LMDTRTAPAELGWTANPSSG  43
```

Query Match   48.7%;
Best Local Similarity   50.0%;
Matches   10;   Conservative   4;   Mismatches   6;   Indels   0;   Gaps   0;
Score 58;   DB 13;   Length 985;
Pred. No. 0.76;

```
ID   Q91735                PRELIMINARY;      PRT;   974 AA.
AC   Q91735;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR) (TCK).
RA   JONES T.L., KARAVANOVA I., MAENO M., ONG R.C., KUNG H.-F., DAAR I.O.;
RT   "Expression of an amphibian homolog of the Eph family of receptor
RT   tyrosine kinases is developmentally regulated.";
RL   Oncogene 10:1111-1117(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC       ROLE IN THE DEVELOPMENT OR FUNCTION OF THE CENTRAL NERVOUS SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED, IT DECREASES AT MID
CC       BLASTULA TRANSITION AND REAPPEARS AT LATE NEURULATION. EXPRESSED
CC       AT HIGHER LEVELS IN THE ANTERIOR AND DORSAL REGIONS OF EMBRYONIC
CC       STAGES 16, 24 AND 37. IN ADULT IT APPEARS TO BE UBIQUITOUSLY
CC       EXPRESSED WITH HIGHER EXPRESSION IN BRAIN AND OVARY. EXPRESSION IN
CC       THE BRAIN, BRACHIAL ARCHES, TRIGEMINAL FACIAL GANGLION, AND THE
CC       RETINA OF SWIMMING TADPOLE STAGE OF DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U14164; AAA74888.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; FALSE_NEG.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT   SIGNAL        1     19      POTENTIAL.
FT   CHAIN        20    985      EPHRIN TYPE-B RECEPTOR XEK.
FT   DOMAIN        6     12      POLY-LEU.
FT   DOMAIN       20    542      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    543    563      POTENTIAL.
FT   DOMAIN      564    985      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       56    105      IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      185    321      CYS-RICH.
FT   DOMAIN      322    431      FIBRONECTIN TYPE-III.
FT   DOMAIN      432    529      FIBRONECTIN TYPE-III.
FT   DOMAIN      620    883      PROTEIN KINASE.
FT   NP_BIND     626    634      ATP (BY SIMILARITY).
FT   DISULFID     63     98      BY SIMILARITY.
FT   BINDING     652    652      ATP (BY SIMILARITY).
FT   ACT_SITE    745    745      BY SIMILARITY.
FT   MOD_RES     779    779      PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    336    336      POTENTIAL.
FT   CARBOHYD    428    428      POTENTIAL.
FT   CARBOHYD    482    482      POTENTIAL.
SQ   SEQUENCE   985 AA;  110104 MW;  2416A8F9 CRC32;
```

RESULT 7

```
Qy    3  LLDSKTIQGELGWISYPSHG  22
         |:|::| || ||:|| |
Db   21  LMDTKWTTSELAWVAYPDSG  40
```

Query Match   47.9%;
Best Local Similarity   45.0%;
Matches   9;   Conservative   4;   Mismatches   7;   Indels   0;   Gaps   0;
Score 57;   DB 13;   Length 974;
Pred. No. 1.1;

```
ID   O43477                PRELIMINARY;      PRT;  1055 AA.
AC   O43477;
DE   KINASE RECEPTOR) (TCK).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96068901.
RA   SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;
RT   "Novel members of the eph receptor tyrosine kinase subfamily expressed
RT   during Xenopus development.";
RL   Oncogene 11:1745-1752(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO IN PRE-SOMITIC
CC       MESODERM, CAUDAL SOMITES, MIDBRAIN, AND CEMENT GLAND. MOST
CC       ABUNDANT IN ADULT BRAIN, EYE, HEART, LUNG AND OVARY. LOWER LEVELS
CC       IN INTESTINE, KIDNEY, OVIDUCT AND PHARYNX.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L43620; AAA93526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL        1     16      POTENTIAL.
FT   CHAIN        17    974      EPHRIN TYPE-B RECEPTOR.
FT   DOMAIN       17    534      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    535    555      POTENTIAL.
FT   DOMAIN      556    974      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       53    102      IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      178    315      CYS-RICH.
FT   DOMAIN      316    423      FIBRONECTIN TYPE-III.
FT   DOMAIN      424    520      FIBRONECTIN TYPE-III.
FT   DOMAIN      609    872      PROTEIN KINASE.
FT   NP_BIND     615    623      ATP (BY SIMILARITY).
FT   DISULFID     60     95      BY SIMILARITY.
FT   BINDING     641    641      ATP (BY SIMILARITY).
FT   ACT_SITE    734    734      BY SIMILARITY.
FT   MOD_RES     768    768      PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    330    330      POTENTIAL.
FT   CARBOHYD    420    420      POTENTIAL.
SQ   SEQUENCE   974 AA;  108263 MW;  BD419F6D CRC32;
```

```
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTEIN-TYROSINE KINASE EPHB2V.
GN   EPHB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 9154673.
RA   IKEGAKI N., TANG X.X., LIU X.G., BIEGEL J.A., ALLEN C., YOSHIOKA A.,
RA   SULMAN E.P., BRODEUR G.M., PLEASURE D.E.;
RT   "Molecular characterization and chromosomal localization of DRT
RT   (EPHT3): a developmentally regulated human protein-tyrosine kinase
RT   gene of the EPH family.";
RL   Hum. Mol. Genet. 4:2033-2045(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   TANG X.X., PLEASURE D.E., IKEGAKI N.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF025304; AAB94602.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PRINTS; PR00014; FNTIPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Tyrosine-protein kinase.
SQ   SEQUENCE   1055 AA;   117492 MW;   EAF28155 CRC32;

Query Match             47.1%;
Best Local Similarity   50.0%;
Matches   10;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;
          Score 56;  DB 4;  Length 1055;
          Pred. No. 1.7;

Qy    3  LLDSKTIQGELGWISYPSHG  22
         | |:|| ||||:||  : ||
Db   23  LMDSTATAELGWMVHPPSG  42
```

```
RESULT 8
ID   O43569
AC   O43569;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1C
DE   (EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1D).
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STEIN E., SCHOECKLMANN H.O., DANIEL T.O.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RC   TISSUE=KIDNEY;
DR   EMBL; AF037333; AAB94627.1; -.
DR   EMBL; AF037334; AAB94628.1; -.
DR   HSSP; P08631; 2HCK.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 2.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTIPEIII.
DR   PRINTS; PR00109; TYRKINASE.
     PRELIMINARY;   PRT;   943 AA.
SQ   SEQUENCE   943 AA;   105291 MW;   C4C78A21 CRC32;

Query Match             46.2%;
Best Local Similarity   45.0%;
Matches    9;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;
          Score 55;  DB 4;  Length 943;
          Pred. No. 2.1;

Qy    3  LLDSKTIQGELGWISYPSHG  22
         |:||| |||| : ||
Db   22  LMDTRTATAELGWTANPASG  41
```

```
RESULT 9
ID   O95142
AC   O95142;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1.
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   STEIN E., HUYNH-DO U., LANE A., CERRETTI D.P., DANIEL T.O.;
RT   "Nck recruitment to Eph receptor, EphB1/ELK, couples ligand activation
RT   to c-Jun kinase.";
RL   J. Biol. Chem. 0:0-0(1997).
DR   EMBL; AF037331; AAD02030.1; -.
DR   HSSP; P05523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Receptor; Kinase.
     PRELIMINARY;   PRT;   984 AA.
SQ   SEQUENCE   984 AA;   110041 MW;   DF5C8ED3 CRC32;

Query Match             46.2%;
Best Local Similarity   45.0%;
Matches    9;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;
          Score 55;  DB 4;  Length 984;
          Pred. No. 2.2;

Qy    3  LLDSKTIQGELGWISYPSHG  22
         |:||| |||| : ||
Db   22  LMDTRTATAELGWTANPASG  41
```

```
RESULT 10
ID   Q07494
AC   Q07494;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CEK6 PROTEIN (FRAGMENT).
GN   CEK6.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93288394.
RA   SAJADI F.G., PASQUALE E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
DR   EMBL; Z19110; CAA79526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00523; 2PTK.
DR   PFAM; PF00536; SAM; 1.
     PRELIMINARY;   PRT;   952 AA.
```

DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
FT   NON_TER   1   1
SQ   SEQUENCE   952 AA; 105994 MW; D939D7D0 CRC32;

Qy   3   LLDSKTIQGELGWISYPSHG   22
             |:|:|| |||||||
Db   3   LMDTRTATAELGWTANPPSG   22

Query Match   44.5%; Score 53; DB 13; Length 952;
Best Local Similarity 45.0%; Pred. No. 4.5;
Matches   9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

RESULT 11
ID   O15197
AC   O15197;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-FAMILY RECEPTOR PROTEIN PRECURSOR (EC 2.7.1.112) (HEP).
GN   EPHB6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97350806.
RA   MATSUOKA H., IWATA N., ITO M., SHIMOYAMA M., NAGATA A., CHIHARA K.,
RA   TAKAI S., MATSUI T.;
RT   "Expression of a kinase-defective Eph-like receptor in the normal
RT   human brain.";
RL   Biochem. Biophys. Res. Commun. 235:487-492(1997).
RN   [2]
RP   SEQUENCE OF 1-804 FROM N.A.
RA   TAUDIEN S., ROSENTHAL A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: KINASE-DEFECTIVE RECEPTOR FOR MEMBERS OF THE EPHRIN-B
CC   FAMILY.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN THE BRAIN AND PANCREAS,
CC   AND WEAK EXPRESSION IN OTHER TISSUES, SUCH AS THE HEART, PLACENTA,
CC   LUNG, LIVER, SKELETAL MUSCLE AND KIDNEY.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; D83492; BAA21560.1; -.
DR   EMBL; AF107256; AAD03058.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602757; -.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; FALSE_NEG.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; FALSE_NEG.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; ATP-binding; Phosphorylation; Receptor; Transmembrane;
KW   Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   16
FT   CHAIN   17   1006   EPH-FAMILY RECEPTOR PROTEIN.
FT   DOMAIN   17   579   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   580   600   POTENTIAL.
FT   DOMAIN   601   1006   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   53   106   IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN   151   161   POLY-SER.
FT   DOMAIN   204   351   CYS-RICH.
FT   DOMAIN   352   468   FIBRONECTIN TYPE-III.
FT   DOMAIN   404   407   POLY-GLY.

FT   DOMAIN   469   566   FIBRONECTIN TYPE-III.
FT   DOMAIN   655   904   PROTEIN KINASE-LIKE.
FT   DOMAIN   866   869   POLY-PRO.
FT   NP_BIND   661   669   ATP (BY SIMILARITY).
FT   DISULFID   60   99   BY SIMILARITY.
FT   CARBOHYD   465   465   POTENTIAL.
SQ   SEQUENCE   1006 AA; 109272 MW; B7C0D81E CRC32;

Qy   3   LLDSKTIQGELGWISYPSHG   22
             |||:||||||||:||
Db   21   LLDDTGETSEIGWLTYPGG   40

Query Match   43.7%; Score 52; DB 4; Length 1006;
Best Local Similarity 45.0%; Pred. No. 6.8;
Matches   9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

RESULT 12
ID   O08644
AC   O08644;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH/ELK RECEPTOR-LIKE PROTEIN PRECURSOR (MEP).
GN   CEKL OR MEP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C X 129 F2; TISSUE=THYMUS;
RX   MEDLINE; 96358627.
RA   GURNIAK C.B., BERG L.J.;
RT   "A new member of the Eph family of receptors that lacks protein
RT   tyrosine kinase activity.";
RL   Oncogene 13:777-786(1996).
CC   -!- FUNCTION: KINASE-DEFECTIVE RECEPTOR FOR MEMBERS OF THE EPHRIN-B
CC   FAMILY.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE TWO SHORTER
CC   ISOFORMS ARE SECRETED.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS ARE PRODUCED BY
CC   ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGH LEVEL IN THYMUS, AND BRAIN. VERY LOW
CC   LEVELS OF EXPRESSION IN KIDNEY, LUNG, LIVER, BONE MARROW, SKELETAL
CC   MUSCLE, SPLEEN FROM 2 WEEK OLD AND ADULT MICE, HEART, TESTES AND
CC   EMBRIONIC STEM CELLS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L77867; AAB51430.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:108444; Cekl.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; FALSE_NEG.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; FALSE_NEG.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; ATP-binding; Phosphorylation; Receptor; Transmembrane;
KW   Glycoprotein; Signal; Immunoglobulin domain; Alternative splicing.
FT   SIGNAL   1   32
FT   CHAIN   33   1014   EPH/ELK RECEPTOR-LIKE PROTEIN.
FT   DOMAIN   33   591   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   592   612   POTENTIAL.
FT   DOMAIN   613   1014   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   69   122   IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN   168   173   POLY-SER.
FT   DOMAIN   214   361   CYS-RICH.
FT   DOMAIN   362   476   FIBRONECTIN TYPE-III.

```
FT   DOMAIN         477    574       FIBRONECTIN TYPE-III.
FT   DOMAIN         663    912       PROTEIN KINASE-LIKE.
FT   DOMAIN         874    877       POLY-PRO.
FT   DOMAIN         328    328       POLY-PRO.
FT   VARSPLIC       329   1014       MISSING (IN CDNA 2).
FT   VARSPLIC       583    583       P -> PCPALPLFEHSRPEVC (IN CDNA 2).
FT   VARSPLIC       583    583       P -> S (IN CDNA 3).
FT   VARSPLIC       584   1014       MISSING (IN CDNA 3).
FT   NP_BIND        669    677       ATP (BY SIMILARITY).
FT   DISULFID        76    115       BY SIMILARITY.
FT   CARBOHYD       473    473       POTENTIAL.
SQ   SEQUENCE      1014 AA;  110076 MW;  F372C039 CRC32;

Qy          3  LLDSKTIQGELGWISPSHG  22
               |||: ||||:||:||
Db         37  LLDTTGETSEIGWLTYPPGG  56

Query Match
Best Local Similarity   43.7%;
Matches     9;   Conservative    4;   Mismatches    7;   Indels    0;   Gaps    0;


RESULT  13
ID   Q50031       PRELIMINARY;      PRT;    366 AA.
AC   Q50031;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE   U2266F.
OS   Mycobacterium leprae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   SMITH D.R.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   ROBISON K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U15182; AAA62976.1; -.
SQ   SEQUENCE      366 AA;  39212 MW;  6C314913 CRC32;

Qy          5  DSKTIQGELGWISYPSHG  22
               |:|| |:||||| |
Db        226  DQRIPHEVGWIGGPGEG  243

Query Match
Best Local Similarity   50.0%;   Pred. No. 3.1;   Score 51;   DB 2;   Length 366;
Matches     9;   Conservative    2;   Mismatches    7;   Indels    0;   Gaps    0;


RESULT  14
ID   O53825       PRELIMINARY;      PRT;    368 AA.
AC   O53825;
DT   01-JUN-1998 (TREMBLrel. 06, Created)
DT   01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT   01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 39.4 KD PROTEIN.
GN   MTV043.03C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RA   HAMLIN N., CHURCHER C.M.;
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RA   COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE; 96181548.
RA   PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA   BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA   COLE S.T.;
RT   "An integrated map of the genome of the tubercle bacillus,
RT   Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RT   leprae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
DR   EMBL; AL022004; CAA17617.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE      368 AA;  39447 MW;  61502D73 CRC32;

Qy          5  DSKTIQGELGWISYPSHG  22
               |:|| |:||||| |
Db        230  DERTIPHEVGWIGGPGAG  247

Query Match
Best Local Similarity   41.2%;   Pred. No. 6.5;   Score 49;   DB 2;   Length 368;
Matches     9;   Conservative    2;   Mismatches    7;   Indels    0;   Gaps    0;


RESULT  15
ID   Q17901       PRELIMINARY;      PRT;    1763 AA.
AC   Q17901; Q20887;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT   01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE   C10C5.6A PROTEIN.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MATTHEWS P.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   MATTHEWS P.;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   MATTHEWS P.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z69662; CAA93506.1; -.
DR   EMBL; Z68214; CAA92448.1; -.
DR   EMBL; Z68214; CAA92448.1; JOINED.
DR   EMBL; Z69662; CAA92448.1; JOINED.
SQ   SEQUENCE     1763 AA;  201072 MW;  EEDA34C5 CRC32;

Qy          5  DSKTIQGELGWISYPSHG  22
               |:|| |:||||| |
Db        247  DERTIPHEVGWIGGPGAG

Query Match
Best Local Similarity   40.3%;   Pred. No. 55;   Score 48;   DB 5;   Length 1763;
                        53.3%;
```

Matches   8;   Conservative   3;   Mismatches   4;   Indels   0;   Gaps   0;

Qy   2   NLLDSKTIQGELGWI 16
            || : :|: ||| ||
Db   359   NLAERRTVLGELNWI 373

Search completed: May 9, 2000, 22:31:10
Job time: 2375 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 22:32:06 ; Search time 49.21 seconds
(without alignments)
13.615 Million cell updates/sec

Title: US-09-104-340-2
Perfect score: 119
Sequence: 119
1 VNLLDSRTIQGELGWISYPSHG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 3045497 3 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	119	100.0	983	1	EPA3_CHICK	P29318 gallus gall
2	119	100.0	983	1	EPA3_HUMAN	P29320 homo sapien
3	119	100.0	983	1	EPA3_MOUSE	P29319 mus musculu
4	119	100.0	984	1	EPA3_RAT	O08680 rattus norv
5	93	78.2	1013	1	EPA5_CHICK	P54755 gallus gall
6	89	74.8	877	1	EPA5_MOUSE	O60629 mus musculu
7	89	74.8	1004	1	EPA8_MOUSE	O09127 mus musculu
8	89	74.8	1037	1	EPA5_HUMAN	P54756 homo sapien
9	87	73.1	1005	1	EPA5_RAT	P54757 rattus norv
10	79	66.4	986	1	EPA4_CHICK	Q07496 gallus gall
11	79	66.4	986	1	EPA4_HUMAN	P54764 homo sapien
12	79	66.4	986	1	EPA4_MOUSE	Q03137 mus musculu
13	73	61.3	948	1	EPA6_RAT	P54758 rattus norv
14	72	60.5	1035	1	EPA6_MOUSE	O62413 mus musculu
15	61	51.3	998	1	EPA7_HUMAN	Q15375 homo sapien
16	59	49.6	998	1	EPA7_MOUSE	O61772 mus musculu
17	59	49.6	998	1	EPA7_RAT	P54759 rattus norv
18	58	48.7	977	1	EPA2_MOUSE	Q03145 mus musculu
19	57	47.9	968	1	SYA_HUMAN	P49588 homo sapien
20	56	47.1	976	1	EPA1_HUMAN	P21709 homo sapien
21	56	47.1	986	1	EPA2_HUMAN	P29323 homo sapien
22	56	47.1	987	1	EPB2_COTJA	Q90344 coturnix co
23	56	47.1	988	1	EPB2_HUMAN	P28693 homo sapien
24	56	47.1	993	1	EPB2_MOUSE	P54763 mus musculu
25	55	46.2	984	1	EPB1_HUMAN	P54762 homo sapien
26	55	46.2	984	1	EPB1_RAT	P09759 rattus norv
27	55	46.2	993	1	EPB3_MOUSE	P54754 mus musculu
28	55	46.2	993	1	EPB3_HUMAN	P54753 homo sapien
29	53	44.5	976	1	EPA2_HUMAN	P29317 homo sapien
30	52	43.7	1006	1	EPB6_HUMAN	O15197 homo sapien
31	48	40.3	1042	1	EPB5_CHICK	Q07497 gallus gall
32	47	39.5	309	1	YOHG_ECOLI	Q46850 escherichia
33	47	39.5	967	1	SYA_BOMMO	P21894 bombyx mori
34	46	38.7	746	1	CHS6_YEAST	P40955 saccharomyc
35	45	37.8	619	1	SELB_HAEIN	P43927 haemophilus
36	45	37.8	1577	1	MYSH_ACACA	P47808 acanthamoeb
37	45	37.8	124	1	PA2_OPHHA	P80966 ophiophagus
38	43	36.1	309	1	WAPP_COCIM	P42783 coccidioide
39	43	36.1	367	1	YMP4_STRCO	P43167 streptomyce
40	43	36.1	551	1	FIKN_AZOCA	P98056 azorhizobiu
41	43	36.1	987	1	EPB4_MOUSE	P54761 mus musculu
42	42.5	35.7	331	1	YIAH_ECOLI	P37669 escherichia
43	42.5	35.7	392	1	DCAM_LEIDO	Q25264 leishmania
44	42.5	35.7	2483	1	COA2_HUMAN	O00705 homo sapien
45	42	35.3	125	1	FRDD_MYCTU	Q10763 mycobacteri

ALIGNMENTS

RESULT 1
EPA3_CHICK

ID   EPA3_CHICK   STANDARD;   PRT;   983 AA.
AC   P29318;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (CEK4).
GN   EPHA3 OR ETK1 OR CEK4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURING IN THE
CC   BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M68514; AAA48666.1; -.
DR   PIR; B45583; B45583.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.

```
FT   SIGNAL        1     19       BY SIMILARITY.
FT   CHAIN        20    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       20    540       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       541   564       POTENTIAL.
FT   TRANSMEM     541   564       POTENTIAL.
FT   DOMAIN       565   983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       20    320       CYS-RICH.
FT   DOMAIN       321   431       FIBRONECTIN TYPE-III.
FT   DOMAIN       432   528       FIBRONECTIN TYPE-III.
FT   DOMAIN       621   882       PROTEIN KINASE.
FT   NP_BIND      627   635       ATP (BY SIMILARITY).
FT   BINDING      653           ATP (BY SIMILARITY).
FT   ACT_SITE     746           BY SIMILARITY.
FT   MOD_RES      779   779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     231   231       POTENTIAL.
FT   CARBOHYD     336   336       POTENTIAL.
FT   CARBOHYD     390   390       POTENTIAL.
FT   CARBOHYD     403   403       POTENTIAL.
FT   CARBOHYD     492   492       POTENTIAL.
SQ   SEQUENCE   983 AA;  109910 MW;  E8895F0BDF77651E CRC64;

Qy            1  VNLLDSKTIQGELGWISYPSHG  22
                 ||||||||||||||||||||||
Db           29  VNLLDSKTIQGELGWISYPSHG  50

Query Match              100.0%;   Score 119;  DB 1;  Length 983;
Best Local Similarity    100.0%;   Pred. No. 1e-10;
Matches   22;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT 2

ID   EPA3_HUMAN     STANDARD;      PRT;   983 AA.
AC   P29320;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (HEK).
GN   EPHA3 OR ETK1 OR ETK OR HEK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE OF 21-29 AND 840-860, AND CHARACTERIZATION.
RX   MEDLINE; 92147681.
RA   Boyd A.W., Ward L.D., Wicks I.P., Simpson R.J., Salvaris E., Wilks A.,
RA   Welch K., Loudovaris M., Rockman S., Busmanis I.;
RL   "Isolation and characterization of a novel receptor-type protein
RL   tyrosine kinase (hek) from a human pre-B cell line.";
RL   J. Biol. Chem. 267:3262-3267(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92179233.
RA   Wicks I.P., Wilkinson D., Salvaris E., Boyd A.W.;
RL   "Molecular cloning of HEK, the gene encoding a receptor tyrosine
RL   kinase expressed by human lymphoid tumor cell lines.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1611-1615(1992).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5. COULD PLAY A ROLE IN LYMPHOID
CC       FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO LYMPHOID TUMOR CELL LINES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; M83941; AAA58633.1; -.
DR   EMBL; A28003; CAA01906.1; -.
DR   PIR; A38224; A38224.
DR   HSSP; P00523; 2PTK.
DR   MIM; 179611; -.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM_1; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     542   565       POTENTIAL.
FT   DOMAIN       566   983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    320       CYS-RICH.
FT   DOMAIN       321   431       FIBRONECTIN TYPE-III.
FT   DOMAIN       432   528       FIBRONECTIN TYPE-III.
FT   DOMAIN       621   882       PROTEIN KINASE.
FT   NP_BIND      627   635       ATP (BY SIMILARITY).
FT   BINDING      653           ATP (BY SIMILARITY).
FT   ACT_SITE     746           BY SIMILARITY.
FT   MOD_RES      779   779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     337   337       POTENTIAL.
FT   CARBOHYD     391   391       POTENTIAL.
FT   CARBOHYD     404   404       POTENTIAL.
FT   CARBOHYD     493   493       POTENTIAL.
FT   CONFLICT     507   507       F -> L (IN CAA01906).
FT   CONFLICT     724   724       V -> L (IN CAA01906).
SQ   SEQUENCE   983 AA;  110086 MW;  B8D900FA80FF5121 CRC64;

Qy            1  VNLLDSKTIQGELGWISYPSHG  22
                 ||||||||||||||||||||||
Db           30  VNLLDSKTIQGELGWISYPSHG  51

Query Match              100.0%;   Score 119;  DB 1;  Length 983;
Best Local Similarity    100.0%;   Pred. No. 1e-10;
Matches   22;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT 3

ID   EPA3_MOUSE     STANDARD;      PRT;   983 AA.
AC   P29319;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (MEK4).
GN   EPHA3 OR ETK1 OR MEK4 OR TYRO4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR X SWISS WEBSTER; TISSUE=EMBRYO;
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
```

Left column:

```
RT
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SECRETED VARIANT WITHOUT THE TRANSMEMBRANE
CC       REGION AND THE KINASE DOMAIN IS PRODUCED BY ALTERNATIVE SPLICING
CC       OF THE SAME GENE.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M68513; AAA39521.1; -.
DR   EMBL; M68515; AAA39522.1; ALT_SEQ.
DR   PIR; A45583; A45583.
DR   HSSP; P16109; 1FSB.
DR   MGD; MGI:99612; EPHA3.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; SAM; 1.
DR   PFAM; PF00536; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    541    564       POTENTIAL.
FT   DOMAIN      565    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    320       CYS-RICH.
FT   DOMAIN      321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN      432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    231    231       POTENTIAL.
FT   CARBOHYD    336    336       POTENTIAL.
FT   CARBOHYD    390    390       POTENTIAL.
FT   CARBOHYD    403    403       POTENTIAL.
FT   CARBOHYD    492    492       POTENTIAL.
FT   CARBOHYD    530    530       POTENTIAL.
FT   VARSPLIC    983    983       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   983 AA;  109955 MW;  BE44A6655D8107A2 CRC64;
```

Right column:

```
Db   30   VNLLDSKTIQGELGWISYPSHG 51

RESULT     4
EPA3_RAT
ID   EPA3_RAT       STANDARD;      PRT;   984 AA.
AC   O08680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR REK4).
GN   EPHA3 OR REK4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE; 98120505.
RA   Li Y.Y., McKiernan C.F., Feldman A.M.;
RT   "IL-1 beta alters the expression of the receptor tyrosine kinase gene
RT   r-EphA3 in neonatal rat cardiomyocytes.";
RL   Am. J. Physiol. 274:H331-H341(1998).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U69278; AAC06273.1; -.
DR   HSSP; P16109; 1FSB.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; SAM; 1.
DR   PFAM; PF00536; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    984       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    542    566       POTENTIAL.
FT   DOMAIN      566    984       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    321       CYS-RICH.
FT   DOMAIN      322    432       FIBRONECTIN TYPE-III.
FT   DOMAIN      433    529       FIBRONECTIN TYPE-III.
FT   DOMAIN      622    883       PROTEIN KINASE.
FT   NP_BIND     628    636       ATP (BY SIMILARITY).
FT   BINDING     654    654       ATP (BY SIMILARITY).
FT   ACT_SITE    747    747       BY SIMILARITY.
FT   MOD_RES     780    780       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    232    232       POTENTIAL.
FT   CARBOHYD    337    337       POTENTIAL.
FT   CARBOHYD    391    391       POTENTIAL.
```

Query match (bottom of left column):

```
Qy

Query Match          100.0%;     Score 119;   DB 1;   Length 983;
Best Local Similarity 100.0%;     Pred. No. 1e-10;
Matches    22;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

  1 VNLLDSKTIQGELGWISYPSHG 22
    |||||||||||||||||||||||
```

```
FT   CARBOHYD       404    404       POTENTIAL.
FT   CARBOHYD       493    493       POTENTIAL.
SQ   SEQUENCE       984 AA;  110227 MW; F170C49312F7A0AB CRC64;

Qy   Query Match
     Best Local Similarity   100.0%;  Score 119; DB 1; Length 984;
     Matches  22; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
       |||||||||||||||||||||||
Db  30 VNLLDSKTIQGELGWISYPSHG 51

RESULT 5
ID   EPA5_CHICK      STANDARD;      PRT;   1013 AA.
AC   P54755;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR CEK7).
GN   EPHA5 OR CEK7.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BODY WALL;
RX   MEDLINE; 95047429.
RA   Siever D.A., Verderame M.F.;
RT   "Identification of a complete Cek7 receptor protein tyrosine kinase
RT   coding sequence and cDNAs of alternatively spliced transcripts.";
RL   Gene 148:219-226(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```
```
DR   EMBL; U03910; AAB60613.1; -.
DR   EMBL; U03910; AAB60614.1; -.
DR   EMBL; U03910; AAB60612.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL           1     31       POTENTIAL.
```

```
FT   CHAIN           32   1013       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN          32    549       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM       550    570       POTENTIAL.
FT   DOMAIN         571   1013       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         651    912       PROTEIN KINASE.
FT   NP_BIND        657    665       ATP (BY SIMILARITY).
FT   BINDING        683    683       ATP (BY SIMILARITY).
FT   ACT_SITE       776    776       BY SIMILARITY.
FT   CARBOHYD       240    240       POTENTIAL.
FT   CARBOHYD       275    275       POTENTIAL.
FT   CARBOHYD       345    345       POTENTIAL.
FT   CARBOHYD       399    399       POTENTIAL.
FT   CARBOHYD       412    412       POTENTIAL.
FT   CARBOHYD       437    437       POTENTIAL.
FT   VARSPLIC       280    280       MISSING (IN ISOFORM 2).
FT   VARSPLIC       573    595       SCDDHGCGWASSLRAVAYPPSLIW -> R (IN ISOFORM
FT                                   1 AND ISOFORM 2).
SQ   SEQUENCE      1013 AA;  112245 MW; AC36FD6AEBF38382 CRC64;

Qy   Query Match
     Best Local Similarity   78.2%;  Score 93; DB 1; Length 1013;
     Matches  15; Conservative  5; Mismatches  2; Indels  0; Gaps  0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
       ||||||::|:|:||||||:||:|
Db  37 VNLLDSRTVMGDLGWIAYPKNG 58

RESULT 6
ID   EPA5_MOUSE      STANDARD;      PRT;    877 AA.
AC   Q60629;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (BRAIN-SPECIFIC KINASE)
DE   (CEK-7).
GN   EPHA5 OR EHK1 OR CEK7 OR BSK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 94194581.
RA   Zhou R., Copeland T.D., Kromer L.F., Schulz N.T.;
RT   "Isolation and characterization of Bsk, a growth factor receptor-like
RT   tyrosine kinase associated with the limbic system.";
RL   J. Neurosci. Res. 37:129-143(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
```
```
DR   EMBL; U07357; AAA17038.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:99654; Epha5.
DR   PRINTS; PR00014; FNTYPEIII.
```

RESULT 6 (continued)

```
Qy    1  VNLLDSKTIQGELGWISYPSHG 22
         |||||:|:|:|:|||||:| :|
Db   63  VNLLDSRTVMGDLGWIAFPKNG 84

Query Match             74.8%;    Score 89; DB 1; Length 877;
Best Local Similarity   63.6%;    Pred. No. 4.1e-06;
Matches 14;  Conservative 6;  Mismatches 2;  Indels 0;  Gaps 0;
```

```
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL         1      26       BY SIMILARITY.
FT   CHAIN         27     877       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN        27     412       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     413     433       POTENTIAL.
FT   DOMAIN       434     877       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       305     399       FIBRONECTIN TYPE-III.
FT   DOMAIN       514     775       PROTEIN KINASE.
FT   NP_BIND      520     528       ATP (BY SIMILARITY).
FT   BINDING      546     546       ATP (BY SIMILARITY).
FT   ACT_SITE     639     639       BY SIMILARITY.
FT   CARBOHYD     266     266       POTENTIAL.
FT   CARBOHYD     301     301       POTENTIAL.
FT   SEQUENCE    877 AA;  97115 MW;  54AD2DC86A178214 CRC64;
```

RESULT 7

```
ID   EPA8_MOUSE    STANDARD;      PRT;   1004 AA.
AC   O09127;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 8 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EEK) (EPH-AND ELK-RELATED KINASE).
GN   EPHA8 OR EEK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97178845.
RA   Park S., Sanchez M.P.;
RT   "The Eek receptor, a member of the Eph family of tyrosine protein
RT   kinases, can be activated by three different Eph family ligands.";
RL   Oncogene 14:533-542(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY.
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; U72207; AAB39218.1; -.
DR   HSSP; P11362; 1FGI.
DR   MGD; MGI:1109378; EPHA8.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL         1       ?       POTENTIAL.
FT   CHAIN          ?    1004       EPHRIN TYPE-A RECEPTOR 8.
FT   DOMAIN         ?     541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     542     562       POTENTIAL.
FT   DOMAIN       563    1004       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND      640     648       ATP (BY SIMILARITY).
FT   BINDING      666     666       ATP (BY SIMILARITY).
FT   ACT_SITE     759     759       BY SIMILARITY.
FT   MOD_RES      792     792       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     339     339       POTENTIAL.
FT   CARBOHYD     406     406       POTENTIAL.
FT   CARBOHYD     431     431       POTENTIAL.
FT   SEQUENCE   1004 AA;  110764 MW;  8530E800290FE502 CRC64;
```

RESULT 8

```
Qy    1  VNLLDSKTIQGELGWISYPSHG 22
         |||||:|:| || :||:||:|
Db   31  VNLLDTSTIHGDWGWLTYPAHG 52

Query Match             74.8%;    Score 89; DB 1; Length 1004;
Best Local Similarity   63.6%;    Pred. No. 4.7e-06;
Matches 14;  Conservative 5;  Mismatches 3;  Indels 0;  Gaps 0;
```

```
ID   EPA5_HUMAN    STANDARD;      PRT;   1037 AA.
AC   P54756;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (RECEPTOR PROTEIN-
DE   TYROSINE KINASE HEK7).
GN   EPHA5 OR EHK1 OR HEK7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Mischer G.C.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 25-1037 FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
```

CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC   ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC   SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR TYROSINE KINASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X95425; CAA64700.1; -.
DR   EMBL; L36644; AAA74245.1; -.
DR   HSSP; P00523; 2PTK.
MIM; 600004; -.
PRINTS; PR00014; FNTYPEIII.
PRINTS; PR00109; TYRKINASE.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
PROSITE; PS01186; EGF_2; UNKNOWN_1.
PFAM; PF00041; fn3; 2.
PFAM; PF00069; pkinase; 1.
PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL   1   24   POTENTIAL.
FT   CHAIN   25   1037   EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN   25   573   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   574   594   POTENTIAL.
FT   DOMAIN   595   1037   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   675   936   PROTEIN KINASE.
FT   NP_BIND   681   689   ATP (BY SIMILARITY).
FT   BINDING   707   707   ATP (BY SIMILARITY).
FT   ACT_SITE   800   800   BY SIMILARITY.
FT   CARBOHYD   264   264   POTENTIAL.
FT   CARBOHYD   299   299   POTENTIAL.
FT   CARBOHYD   369   369   POTENTIAL.
FT   CARBOHYD   423   423   POTENTIAL.
FT   CARBOHYD   436   436   POTENTIAL.
FT   CARBOHYD   461   461   POTENTIAL.
FT   VARSPLIC   597   619   SCCECGCGRASSLCAVAHPLIIW -> R (IN ISOFORM
FT   2).
SQ   SEQUENCE   1037 AA; 114784 MW; FC2C46C959AFB699 CRC64;

RESULT 9

ID   EPA5_RAT   STANDARD;   PRT; 1005 AA.
AC   P54757;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (TYROSINE-PROTEIN
DE   KINASE EHK1 OR EHK-1).
GN   EPHA5 OR EHK1 OR EHK-1.

Query Match   74.8%; Score 89; DB 1; Length 1037;
Best Local Similarity   63.6%; Pred. No. 4.9e-06;
Matches 14; Conservative   6; Mismatches   2; Indels   0; Gaps   0;

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||||:|:||||||:||||:| :
Db   61 VNLLDSRTVMGDLGWIAFPKNG 82
```

OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 94067777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 95206467.
RA   Taylor V., Pfarr S., Miescher G.C., Honegger P., Breitschopf H.,
RA   Lassmann H., Pfarr S.;
RT   "Expression and developmental regulation of Ehk-1, a neuronal
RT   Elk-like receptor tyrosine kinase in brain.";
RL   Neuroscience 63:163-178(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC   ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN NEURONS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR TYROSINE KINASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X76689; CAA54357.1; -.
DR   HSSP; P00523; 2PTK.
PRINTS; PR00014; FNTYPEIII.
PRINTS; PR00109; TYRKINASE.
PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
PROSITE; PS01186; EGF_2; UNKNOWN_1.
PFAM; PF00041; fn3; 2.
PFAM; PF00069; pkinase; 1.
PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL   1   ?   POTENTIAL.
FT   CHAIN   ?   1005   EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN   ?   575   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   576   596   POTENTIAL.
FT   DOMAIN   597   1005   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   677   938   PROTEIN KINASE.
FT   NP_BIND   683   691   ATP (BY SIMILARITY).
FT   BINDING   709   709   ATP (BY SIMILARITY).
FT   ACT_SITE   802   802   BY SIMILARITY.
FT   CARBOHYD   266   266   POTENTIAL.
FT   CARBOHYD   301   301   POTENTIAL.
FT   CARBOHYD   371   371   POTENTIAL.
FT   CARBOHYD   425   425   POTENTIAL.
FT   CARBOHYD   438   438   POTENTIAL.
FT   CARBOHYD   463   463   POTENTIAL.
FT   VARSPLIC   10   20   GRRTQGRGGG -> DADGPRAQASWCHARR (IN
FT   SPLICED FORMS).

```
FT   VARSPLIC    306    358    VCRPGFFKASPHSQTCSKCPPHSYTHEEASTSCVCERDYFR
FT                             RESDPPTWACTR -> G (IN SPLICED FORMS).
FT   VARSPLIC    358    470    RPPSAPRNAISNVNETSVFLEWIPPADTGGGKDVSYIILCK
FT                             KCNSHAGVCEECGGHVRLPQQIGLIKNTSVMMADPLAHTNY
FT                             TEIEAVNGVSDLSPGTRQYVSVNVTINQAA -> T (IN
FT                             SPLICED FORMS).
FT   VARSPLIC    597    621    SGSCCGGCGGRASSLCAVAHPSLIW -> R (IN
FT                             SPLICED FORMS).
FT   CONFLICT    170    170    D -> E (IN REF. 2).
FT   CONFLICT    566    566    D -> A (IN REF. 2).
FT   CONFLICT    578    578    G -> A (IN REF. 2).
FT   CONFLICT    669    669    G -> A (IN REF. 2).
FT   CONFLICT    708    708    T -> I (IN REF. 2).
FT   CONFLICT    979    979    T -> I (IN REF. 2).
SQ   SEQUENCE   1005 AA;  111007 MW;  1AED42C99693C574 CRC64;

Qy     1 VNLLDSKTIQGELGWISYPSHG 22
         ||||||:||: |:||||||:| :|
Db    63 VNLLDSRTVLGDLGWIAFPKNG 84

Query Match           73.1%;  Score 87;  DB 1;  Length 1005;
Best Local Similarity 63.6%;  Pred. No. 9.6e-06;
Matches   14;  Conservative   6;  Mismatches   2;  Indels   0;  Gaps   0;

RESULT 10
EPA4_CHICK
ID   EPA4_CHICK     STANDARD;     PRT;    986 AA.
AC   Q07496; Q90772;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR CEK8).
GN   EPHA4 OR CEK8.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPINAL CORD;
RX   MEDLINE; 96404128.
RA   Ohta K., Nakamura M., Hirokawa K., Tanaka S., Iwama A., Suda T.,
RA   Ando M., Tanaka H.;
RT   "The receptor tyrosine kinase, Cek8, is transiently expressed on
RT   subtypes of motoneurons in the spinal cord during development.";
RL   Mech. Dev. 54:59-69(1996).
RN   [2]
RP   SEQUENCE OF 138-986 FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 93288394.
RA   Sajjadi F.G., Pasquale E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; D38174; BAA07373.1; -.
DR   EMBL; Z19059; CAA79509.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; ENTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Transferase; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CONFLICT    138    138       R -> G (IN REF. 2).
FT   CONFLICT    487    487       S -> T (IN REF. 2).
SQ   SEQUENCE   986 AA;  BD88C2A5BD840A0F CRC64;

Qy     1 VNLLDSKTIQGELGWISYPSHG 22
         ||||||:||:||||||:||:| |
Db    31 VTLLDSRVQGELGWIASPLEG 52

Query Match           66.4%;  Score 79;  DB 1;  Length 986;
Best Local Similarity 63.6%;  Pred. No. 0.00017;
Matches   14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

RESULT 11
EPA4_HUMAN
ID   EPA4_HUMAN     STANDARD;     PRT;    986 AA.
AC   P54764;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (RECEPTOR PROTEIN-TYROSINE KINASE HEK8).
GN   EPHA4 OR SEK OR HEK8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
```

```
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattei M.-G.,
RA   Chestier A., Wilkinson D.G., Charnay P.;
RT   "An Eph-related receptor protein tyrosine kinase gene segmentally
RT   expressed in the developing mouse hindbrain.";
RL   Oncogene 7:2499-2506(1992).
RN   [2]
RP   ERRATUM.
RX   MEDLINE; 93205393.
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattei M.-G.,
RA   Chestier A., Wilkinson D.G., Charnay P.;
RL   Oncogene 8:1103-1103(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC       MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PROCESS INVOLVED IN
CC       HINDBRAIN PATTERN FORMATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SHORTER FORM WITH A DELETION OF 49
CC       AMINO ACIDS ALTERING THE CATALYTIC SITE MAY BE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE BRAIN, LOWER LEVELS
CC       IN THE HEART, LUNG AND KIDNEY. IT IS SEGMENTALLY EXPRESSED
CC       IN THE EMBRYONIC BRAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; X65138; CAA46268.1; -.
DR   EMBL; X57241; CAA40517.1; -.
DR   EMBL; S57168; AAB25836.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:98277; EPHA4.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CARBOHYD    423    423       POTENTIAL.
FT   VARSPLIC    783    832       MISSING (IN SHORT ISOFORM).
SEQUENCE    986 AA;  109801 MW;  D16AD8B85668C80E CRC64;
```

```
RESULT 12
EPA4_MOUSE
ID   EPA4_MOUSE   STANDARD;   PRT;   986 AA.
AC   O03137;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (MPK-3).
GN   EPHA4 OR SEK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 93096484.
```

```
Query Match     66.4%;  Score 79;  DB 1;  Length 986;
Best Local Similarity   63.6%;  Pred. No. 0.00017;
Matches   14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

Qy     1  VNLLDSKHIQGELGWISYPSHG 22
          | |||:::||||||||||:|| |
Db    31  VTLLDSRSVQGELGWIASPLEG 52
```

```
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; L36645; AAA74246.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602188; -.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CARBOHYD    545    545       POTENTIAL.
SEQUENCE    986 AA;  109859 MW;  0C39C11522EDDD46F CRC64;
```

RESULT 13

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        | |||:::||||||||:|| |
Db   31 VTLLDSRSVQGELGWIASPLEG 52
```

Query Match   66.4%; Score 79; DB 1; Length 986;
Best Local Similarity   63.68%; Pred. No. 0.00017;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
ID   EPA6_RAT     STANDARD;      PRT;   948 AA.
AC   P54758;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 6 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-2).
GN   EPHA6 OR EHK2 OR EHK-2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 94067777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    948       EPHRIN TYPE-A RECEPTOR 6.
FT   DOMAIN       23    549       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    550    570       POTENTIAL.
FT   DOMAIN      571    948       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      630    943       PROTEIN KINASE.
FT   NP_BIND     636    644       ATP (BY SIMILARITY).
FT   BINDING     662    662       ATP (BY SIMILARITY).
FT   ACT_SITE    797    797       BY SIMILARITY.
FT   CARBOHYD    342    342       POTENTIAL.
FT   CARBOHYD    396    396       POTENTIAL.
FT   CARBOHYD    409    409       POTENTIAL.
SQ   SEQUENCE   948 AA;  106235 MW;  A47DC78EDB2DEF30 CRC64;
```

Query Match   61.3%; Score 73; DB 1; Length 948;
Best Local Similarity   59.18%; Pred. No. 0.0014;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        | |||:||||||||:||:|| :|
Db   22 ...
```

RESULT 14

```
Db   34 VVLLDIDTVMGELGWKIYPLNG 55
```

```
ID   EPA6_MOUSE    STANDARD;     PRT;   1035 AA.
AC   Q62413;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 6 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-2).
GN   EPHA6 OR EHK2 OR EHK-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE; 97047913.
RA   Lee A.M., Navarathan D., Ichimiya S., Greene M.I., Davis J.G.;
RT   "Cloning of m-ehk2 from the murine inner ear, an eph family receptor
RT   tyrosine kinase expressed in the developing and adult cochlea.";
RL   DNA Cell Biol. 15:817-825(1996).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U58332; AAB53836.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:108034; EPHA6.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23   1035       EPHRIN TYPE-A RECEPTOR 6.
FT   DOMAIN       23    549       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    550    570       POTENTIAL.
FT   DOMAIN      571   1035       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      630    943       PROTEIN KINASE.
FT   NP_BIND     636    644       ATP (BY SIMILARITY).
FT   BINDING     662    662       ATP (BY SIMILARITY).
FT   ACT_SITE    797    797       BY SIMILARITY.
FT   CARBOHYD    342    342       POTENTIAL.
FT   CARBOHYD    396    396       POTENTIAL.
FT   CARBOHYD    409    409       POTENTIAL.
SQ   SEQUENCE  1035 AA;  116137 MW;  560B26419A5EF74 CRC64;
```

```
RESULT 15

Query Match                    60.58;   Score 72;  DB 1;   Length 1035;
Best Local Similarity  59.18;  Pred. No. 0.0021;
Matches  13;  Conservative  4;  Mismatches  5;  Indels  0;  Gaps  0;

Qy      1 VNLLDSKTIQGELGWISYPSHG 22
          ||||:|| ||||||:| ||||:| :|
Db     34 VVLLDTTVWGELGWKYPLNG 55
```

```
ID  EPA7_HUMAN     STANDARD;      PRT;   998 AA.
AC  Q15375;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE  KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3) (RECEPTOR PROTEIN-
DE  TYROSINE KINASE HEK11).
DE  EPHA7 OR EHK3 OR HEK11.
GN  EPHA7.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=BRAIN;
RX  MEDLINE; 95206782.
RA  Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA  Basu R., Welcher A.A.;
RT  "cDNA cloning and tissue distribution of five human EPH-like receptor
RT  protein-tyrosine kinases.";
RL  Oncogene 10:897-905(1995).
CC  -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC      EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC  -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC      PROTEIN TYROSINE PHOSPHATE.
CC  -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC      DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC  -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC  -----------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----------------------------------------------------------------------
DR  EMBL; L36642; AAA74243.1; -.
DR  HSSP; P00523; 2PTK.
DR  MIM; 602190; -.
DR  PRINTS; PR00014; FNTYPEIII.
DR  PRINTS; PR00109; TYRKINASE.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR  PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR  PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR  PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR  PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR  PFAM; PF00041; fn3; 2.
DR  PFAM; PF00069; pkinase; 1.
DR  PFAM; PF00536; SAM; 1.
DR  PFAM; PF01404; EPH_lbd; 1.
KW  Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW  Receptor; Transmembrane; Glycoprotein; Signal.
FT  SIGNAL        1     24       POTENTIAL.
FT  CHAIN        25    998       EPHRIN TYPE-A RECEPTOR 7.
FT  DOMAIN       25    556       EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM    557    577       POTENTIAL.
FT  DOMAIN      578    998       CYTOPLASMIC (POTENTIAL).
FT  DOMAIN      329    438       FIBRONECTIN TYPE-III.
FT  DOMAIN      439    537       FIBRONECTIN TYPE-III.
FT  DOMAIN      633    894       PROTEIN KINASE.
FT  NP_BIND     639    647       ATP (BY SIMILARITY).
FT  BINDING     665    665       ATP (BY SIMILARITY).
FT  ACT_SITE    758    758       BY SIMILARITY.
FT  MOD_RES     791    791       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT  CARBOHYD    343    343       POTENTIAL.
FT  CARBOHYD    410    410       POTENTIAL.
SQ  SEQUENCE   998 AA;  112096 MW;  479B9CA0D2BB06EB CRC64;
```

```
Query Match                    51.38;   Score 61;  DB 1;   Length 998;
Best Local Similarity  63.68;  Pred. No. 0.11;
Matches  14;  Conservative  1;  Mismatches  7;  Indels  0;  Gaps  0;

Qy      1 VNLLDSKTIQGELGWISYPSHG 22
          |:||||||||||:||:||| :|
Db     33 VLLLDSKAQQTELEWISSPPNG 54
```

```
Search completed: May   9, 2000, 22:32:08
Job time: 2399 sec
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May   9, 2000, 22:27:14 ; Search time 64.29 Seconds
                  (without alignments)
                  20.063 Million cell updates/sec

Title: US-09-104-340-2
Perfect score: 119
Sequence: 1 VNLLDSKTIQGELGWISYPSHG 22

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Database :   PIR_63:*
                  1: pir1:*
                  2: pir2:*
                  3: pir3:*
                  4: pir4:*

Searched: 168808 seqs, 5862743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	119	100.0	983	2	B45583	receptor tyrosine
2	119	100.0	983	2	JC5526	protein-tyrosine k
3	119	100.0	983	2	A45583	receptor tyrosine
4	93	78.2	1013	2	I50615	receptor-type prot
5	89	74.8	877	2	I48967	brain-specific kin
6	89	74.8	991	2	I78843	receptor protein-t
7	87	73.1	893	2	S51603	receptor-like tyro
8	87	73.1	898	2	S47489	receptor-like tyro
9	87	73.1	981	2	S51604	receptor tyrosine
10	87	73.1	1005	2	S49015	receptor-like tyro
11	79	66.4	985	2	I51149	receptor tyrosine
12	79	66.4	986	2	S78059	receptor tyrosine
13	79	66.4	986	2	I78844	protein-tyrosine k
14	73	61.3	948	2	S51605	receptor-like tyro
15	61	51.3	998	2	I18351	receptor protein-t
16	59	49.6	605	2	JC5673	receptor tyrosine
17	59	49.6	610	2	I48612	receptor tyrosine
18	59	49.6	626	2	I48614	developmental kina
19	59	49.6	998	2	JC5672	developmental kina
20	58	48.7	975	2	I48974	receptor-protein t
21	58	48.7	977	2	S49004	tyrosine kinase Mp
22	58	48.7	985	2	I51672	receptor tyrosine
23	57	47.9	985	2	I60107	alanine--tRNA liga
24	56	47.1	970	2	I78842	receptor protein-t
25	56	47.1	984	1	A34076	receptor tyrosine k
26	56	47.1	995	2	A56599	embryo kinase 5 -
27	55	46.2	984	2	A39753	protein-tyrosine k
28	55	46.2	993	2	I48653	mouse developmenta
29	55	46.2	998	2	S37627	protein-tyrosine k
30	53	44.5	952	2	I50612	protein-tyrosine k
31	53	44.5	976	2	A36355	protein-tyrosine k
32	53	43.7	1006	2	JC5526	kinase-defective E
33	49	41.2	368	2	C70809	hypothetical prote
34	48	40.3	1763	2	T19183	hypothetical prote
35	48	40.3	1769	2	T19184	hypothetical prote
36	47	39.5	309	2	C65088	hypothetical prote
37	47	39.5	967	1	SYMYAT	hypothetical prote
38	47	38.7	134	2	S28678	alanine--tRNA liga
39	46	38.7	457	2	G71614	hypothetical prote
40	46	38.7	464	2	S50226	CSD3 protein - yea
41	46	38.7	530	2	T24011	eIF-4A-like DEAD f
42	45	37.8	746	2	B37471	capsid protein - N
43	45	37.8	457	2	I64087	translation elonga
44	44	37.8	1576	1	A23662	myosin I, high mol
45	45	37.0	234	2	A71666	hypothetical prote

ALIGNMENTS

RESULT 1
B45583
receptor tyrosine kinase Cek4 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
A;Accession: B45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse
A;Reference number: A45583; MUID:92031278
A;Note: sequence extracted from NCBI backbone (NCBIN:62405, NCBIP:62411)
A;Status: preliminary
A;Accession: B45583
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68514; NID:g454809; PIDN:AAA48666.1; PID:g211447
C;Superfamily: protein-tyrosine kinase; receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   100.0%; Score 119; DB 2; Length 983;
Best Local Similarity   100.0%; Pred. No. 1.4e-10;
Matches   22; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||||||||||||||||||
Db   29 VNLLDSKTIQGELGWISYPSHG 50

RESULT 2
A38224
protein-tyrosine kinase (EC 2.7.1.112) hek precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C;Accession: A38224; B38224
R;Wicks, I.P.; Wilkinson, D.; Salvaris, E.; Boyd, A.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 1611-1615, 1992
A;Title: Molecular cloning of HEK, the gene encoding a receptor tyrosine kinase expre
A;Reference number: A38224; MUID:92179233
A;Accession: A38224
A;Molecule type: mRNA
A;Residues: 1-983 <WIC>
A;Cross-references: GB:M83941; NID:g183931; PIDN:AAA58633.1; PID:g183932
A;Experimental source: pre-B-cell leukemia cell line LK63
A;Note: sequence extracted from NCBI backbone (NCBIP:86627)
A;Accession: B38224
A;Molecule type: protein
A;Residues: 21-39;810-860 <WI2>
C;Superfamily: protein-tyrosine kinase; receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferas
F;1-20/Domain: signal sequence #status predicted <SIG>

Query Match 100.0%; Score 119; DB 2; Length 983;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1 VNLLDSKTIQGELGWISYPSHG 22
       |||||||||||||||||||||
Db 30 VNLLDSKTIQGELGWISYPSHG 51
```

RESULT 3
A45583
receptor tyrosine kinase Mek4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse an
New Biol. 3, 769-778, 1991
A;Reference number: A45583; MUID:9201278
A;Accession: A45583
A;Molecule type: mRNA
A;Status: Preliminary
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68513; NID:g199119; PIDN:AAA39521.1; PID:g199120
A;Note: sequence extracted from NCBI backbone (NCBIN:62398, NCBIP:62401)
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

RESULT 4
I50615
receptor-type protein-tyrosine kinase Cek7, long splice form - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I50615; I50616; I50614
R;Siever, D.A.; Verderame, M.F.
Gene 148, 219-225, 1994
A;Title: Identification of a complete Cek7 receptor protein tyrosine kinase coding seque
A;Reference number: I50614; MUID:95047429
A;Accession: I50615
A;Molecule type: mRNA
A;Status: translated from GB/EMBL/DDBJ
A;Residues: 1-1013 <SIE>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60612.1; PID:g555618
A;Accession: I50616
A;Molecule type: mRNA
A;Status: translated from GB/EMBL/DDBJ
A;Residues: 1-572, 'R', 596-1013 <SI2>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60613.1; PID:g555619
A;Accession: I50614
A;Molecule type: mRNA
A;Status: translated from GB/EMBL/DDBJ
A;Residues: 1-279, 444-572, 'R', 596-1013 <SI3>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60614.1; PID:g555620
A;Genetics:

F;21-983/Product: protein-tyrosine kinase hek #status experimental <MAT>
F;542-565/Domain: transmembrane #status predicted <TMM>
C;Keywords: protein kinase homology <KIN>
F;611-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;232,337,391,404,493/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Qy  1 VNLLDSKTIQGELGWISYPSHG 22
Db 30 VNLLDSKTIQGELGWISYPSHG 51
```

A;Gene: Cek7
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: alternative splicing; ATP; transmembrane protein
F;649-915/Domain: protein kinase homology <KIN>
F;657-665/Region: protein kinase ATP-binding motif
F;938-1004/Domain: SAM homology <SAM>

Query Match 78.2%; Score 93; DB 2; Length 1013;
Best Local Similarity 68.2%; Pred. No. 1.7e-06;
Matches 15; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy  1 VNLLDSKTIQGELGWISYPSHG 22
Db 37 VNLLDSRTVMGDLGWIAYPKNG 58
```

RESULT 5
I48967
brain-specific kinase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C;Accession: I48967
R;Zhou, R.R.; Copeland, T.D.; Kromer, L.F.; Schulz, N.T.
J. Neurosci. Res. 37, 129-143, 1994
A;Title: Isolation and characterization of Bsk, a growth factor receptor-like tyrosin
A;Reference number: I48967; MUID:94194581
A;Accession: I48967
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-877 <RES>
A;Cross-references: EMBL:U07357; NID:g466369; PIDN:AAA17038.1; PID:g466370
A;Gene: Bsk
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;512-778/Domain: protein kinase homology <KIN>
F;520-528/Region: protein kinase ATP-binding motif
F;801-868/Domain: SAM homology <SAM>

Query Match 74.8%; Score 89; DB 2; Length 877;
Best Local Similarity 63.6%; Pred. No. 6.2e-06;
Matches 14; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

```
Qy  1 VNLLDSKTIQGELGWISYPSHG 22
Db 63 VNLLDSRTVMGDLGWIAFPKNG 84
```

RESULT 6
I78843
receptor protein-tyrosine kinase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999
C;Accession: I78843
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welch
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein
A;Reference number: I58351; MUID:95206782
A;Accession: I78843
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-991 <RES>
A;Cross-references: GB:L36644; NID:g551611; PIDN:AAA74245.1; PID:g551612
A;Gene: HEK7
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Genetics:
F;627-893/Domain: protein kinase homology <KIN>
F;916-982/Domain: SAM homology <SAM>

Query Match 74.8%; Score 89; DB 2; Length 991;

RESULT 7
S51603

```
Qy          1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||:|:|:||||||:|:|
Db         37 VNLLDSRTVMGDLGWIAFPKNG 58
```

Best Local Similarity   63.6%;   Pred. No. 7.1e-06;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S51603
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine fam
A;Reference number: S49015; MUID:94067777
A;Accession: S51603
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-893 <MAI>
A;Cross-references: EMBL:S68028
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;563-829/Domain: protein kinase homology <KIN>
F;571-579/Region: protein kinase ATP-binding motif

RESULT 8
S47489

```
Qy          1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||:|:|:||||||:|:|
Db         63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Best Local Similarity   73.1%;   Score 87;   DB 2;   Length 893;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

receptor tyrosine kinase - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C;Accession: S47489
R;Taylor, V.; Pfarr, S.; Miescher, G.C.; Honegger, P.; Breitschopf, H.; Lassmann, H.; St
submitted to the EMBL Data Library, April 1994
A;Description: Expression and developmental regulation of EHK-1, a neuronal ELK-like rec
A;Reference number: S47489
A;Accession: S47489
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-898 <TAY>
A;Cross-references: EMBL:X78689; NID:g531543; PIDN:CAA55357.1; PID:g531544
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; transmembrane protein
F;568-834/Domain: protein kinase homology <KIN>
F;576-584/Region: protein kinase ATP-binding motif

RESULT 9
S51604

```
Qy          1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||:|:|:||||||:|:|
Db         68 VNLLDSRTVLGDLGWIAFPKNG 89
```

Best Local Similarity   63.6%;   Pred. No. 1.3e-05;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51604
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:94067777
A;Accession: S51604
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-981 <MAI>
A;Cross-references: EMBL:S68029
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;651-917/Domain: protein kinase homology <KIN>
F;659-667/Region: protein kinase ATP-binding motif

RESULT 10
S49015

```
Qy          1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||:|:|:||||||:|:|
Db         63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Best Local Similarity   73.1%;   Score 87;   DB 2;   Length 981;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

receptor tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S49015; S51602
A;Note: the authors translated the codon GAC for residue 170 as Glu
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
A;Reference number: S49015; MUID:94067777
A;Accession: S49015
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1005 <MAI>
A;Cross-references: EMBL:S68024
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Molecule type: mRNA
A;Residues: 1-305; G',359-1005 <MA2>
A;Cross-references: EMBL:S68026
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;675-941/Domain: protein kinase homology <KIN>
F;683-691/Region: protein kinase ATP-binding motif

RESULT 11
I51549

```
Qy          1 VNLLDSKTIQGELGWISYPSHG 22
              ||||||:|:|:||||||:|:|
Db         63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Best Local Similarity   73.1%;   Score 87;   DB 2;   Length 1005;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

receptor tyrosine kinase - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I51549
R;Winning, R.S.; Sargent, T.D.
Mech. Dev. 46, 219-229, 1994
A;Title: Pagliaccio, a member of the Eph family of receptor tyrosine kinase genes, ha

RESULT 12
S78059
protein-tyrosine kinase (EC 2.7.1.112) Eph precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999
C;Accession: S78059; S30505; I58366
R;Charnay, P.
A;Reference number: S78059
A;Molecule type: mRNA
A;Residues: 1-986 <CHA>
A;Cross-references: EMBL:X65138; NID:g54083; PIDN:CAA46268.1; PID:g54084
R;Gilardi-Hebenstreit, P.; Nieto, M.A.; Fraín, M.; Mattei, M.G.; Chestier, A.; Wilkinson
Oncogene 7, 2499-2506, 1992
A;Reference number: S30496; MUID:93096484
A;Accession: S30505
A;Molecule type: mRNA
A;Residues: 1-31, 55-986 <GIL>
A;Cross-references: EMBL:X65138
A;Genetics:
A;Gene: Sek
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferase;
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-986/Product: protein-tyrosine kinase Eph #status predicted <MAT>
F;548-569/Domain: transmembrane #status predicted <TMM>
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;908-974/Domain: SAM homology <SAM>
F;235,340,408,423/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.4%; Score 79; DB 2; Length 986;
Best Local Similarity 63.6%; Pred. No. 0.00026;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
     |||||::|||||||||| | |
Db   31 VTLLDSRSVQGELGWIASPLEG 52

RESULT 13
I78844
receptor protein-tyrosine kinase - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999
C;Accession: I78844
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welcher,

Query Match 66.4%; Score 79; DB 2; Length 986;
Best Local Similarity 63.6%; Pred. No. 0.00026;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
     |||||::|||||||||| | |
Db   31 VTLLDSRSVQGELGWIASPLEG 52

A;Reference number: I51549; MUID:95001564
A;Accession: I51549
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-985 <WIN>
A;Cross-references: GB;L26099; NID:g416402; PIDN:AAA64464.1; PID:g416403
A;Gene: Pag
C;Genetics:
C;Keywords: ATP; transmembrane protein
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
F;18-884/Domain: protein kinase homology <KIN>
F;626-634/Region: protein kinase ATP-binding motif

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
     |||||::|||||||||| | |
Db   31 VTLLDSRSVQGELGWIASPLEG 52

Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein
A;Reference number: I78844
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-986 <RES>
A;Cross-references: GB;L36645; NID:g551613; PIDN:AAA74246.1; PID:g551614
A;Gene: HEK8
C;Genetics:
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
F;619-885/Domain: protein kinase homology <KIN>

RESULT 14
S51605
receptor-like tyrosine kinase Ehk-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51605
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:94067777
A;Molecule type: mRNA
A;Residues: 1-948 <MAI>
A;Status: Preliminary
A;Accession: S51605
A;Cross-references: EMBL:S68030
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;620-936/Domain: protein kinase homology <KIN>
F;636-644/Region: protein kinase ATP-binding motif

Query Match 66.4%; Score 79; DB 2; Length 948;
Best Local Similarity 63.6%; Pred. No. 0.00026;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
     |||||::|||||||||| | |
Db   31 VTLLDSRSVQGELGWIASPLEG 52

RESULT 15
I58351
receptor protein-tyrosine kinase - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C;Accession: I58351
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welch
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein
A;Reference number: I58351; MUID:95206782
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-998 <RES>
A;Cross-references: GB;L36642; NID:g551607; PIDN:AAA74243.1; PID:g551608
A;Gene: HEK11
C;Genetics:
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
F;631-897/Domain: protein kinase homology <KIN>

Query Match 61.3%; Score 73; DB 2; Length 948;
Best Local Similarity 59.1%; Pred. No. 0.0022;
Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
     |||||:|:||||||:|| :|
Db   34 VVLLDTSTVMGELGWKTYPLNG 55

F;639-647/Region: protein kinase ATP-binding motif
F;920-986/Domain: SAM homology <SAM>

Query Match   51.3%;   Score 61;   DB 2;   Length 998;
Best Local Similarity   63.6%;   Pred. No. 0.18;
Matches   14; Conservative   1; Mismatches   7; Indels   0; Gaps   0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
           | /||||| | || ||| | :|
Db   33 VLLLDSKAQQTELEWISSPPNG 54

Search completed: May   9, 2000, 22:27:16
Job time: 4155 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 22:25:58 ; Search time 44.2 Seconds
(without alignments)
7.199 Million cell updates/sec

Title:	US-09-104-340-2
Perfect score:	119
Sequence:	1 VNLLDSKVIQGELGWISYPSHG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 143561 seqs, 1446340 residues

Total number of hits satisfying chosen parameters: 143561

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :   Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	982	2	US-08-673-789-4	Sequence 4, Appli
2	119	100.0	983	1	US-08-167-809-16	Sequence 16, Appl
3	119	100.0	983	1	US-08-167-919A-10	Sequence 10, Appl
4	119	100.0	983	2	US-08-449-645A-21	Sequence 21, Appl
5	119	100.0	983	2	US-08-702-367A-21	Sequence 21, Appl
6	119	100.0	983	4	US-08-715-106-10	Sequence 10, Appl
7	119	100.0	983	4	PCT-US95-04681-21	Sequence 21, Appl
8	119	100.0	877	3	US-08-673-789-2	Sequence 2, Appli
9	89	74.8	877	4	US-08-449-645A-30	Sequence 30, Appl
10	89	74.8	967	2	US-08-702-367A-30	Sequence 30, Appl
11	89	74.8	967	2	US-08-449-645A-13	Sequence 13, Appl
12	89	74.8	991	2	US-08-702-367A-13	Sequence 13, Appl
13	89	74.8	991	4	PCT-US95-04681-13	Sequence 13, Appl
14	87	73.1	991	1	US-08-222-616-36	Sequence 36, Appl
15	87	73.1	928	2	US-08-440-248-2	Sequence 2, Appli
16	87	73.1	928	2	US-08-469-537A-103	Sequence 103, App
17	79	66.4	1005	3	US-08-673-789-3	Sequence 3, Appli
18	79	66.4	986	2	US-08-449-645A-15	Sequence 15, Appl
19	79	66.4	986	2	US-08-702-367A-15	Sequence 15, Appl
20	79	66.4	986	4	PCT-US95-04681-15	Sequence 15, Appl
21	79	66.4	986	1	US-08-222-616-36	Sequence 36, Appl
22	79	66.4	1104	2	PCT-US95-04228-36	Sequence 36, Appl
23	75.5	63.4	1104	2	US-08-449-645A-29	Sequence 29, Appl
24	75.5	63.4	687	2	US-08-702-367A-29	Sequence 29, Appl
25	73	61.3	687	2	US-08-469-537A-101	Sequence 101, App
26	61	51.3	948	2	US-08-449-645A-17	Sequence 17, Appl
27	61	51.3	998	2	US-08-702-367A-17	Sequence 17, Appl
28	61	51.3	998	4	PCT-US95-04681-17	Sequence 17, Appl
29	59	49.6	998	4	PCT-US96-00419-3	Sequence 3, Appli
30	59	49.6	610	4	PCT-US96-00419-5	Sequence 5, Appli
31	59	49.6	626	4	PCT-US96-00419-2	Sequence 2, Appli
32	58	48.7	998	2	US-08-673-789-6	Sequence 6, Appli
33	57	47.9	984	1	US-08-426-236-2	Sequence 2, Appli
34	56	47.1	968	4	US-08-449-645A-11	Sequence 11, Appl
35	56	47.1	970	2	US-08-702-367A-11	Sequence 11, Appl
36	56	47.1	970	4	PCT-US95-04681-11	Sequence 11, Appl
37	56	47.1	970	2	US-08-673-789-9	Sequence 9, Appli
38	56	47.1	984	2	US-08-449-645A-19	Sequence 19, Appl
39	56	47.1	984	2	US-08-702-367A-19	Sequence 19, Appl
40	56	47.1	984	4	PCT-US95-04681-19	Sequence 19, Appl
41	56	47.1	984	1	US-08-162-809-18	Sequence 18, Appl
42	56	47.1	995	2	US-08-673-789-5	Sequence 5, Appli
43	56	47.1	995	2	US-08-162-809-12	Sequence 12, Appl
44	56	46.2	1011	2	US-08-449-645A-20	Sequence 20, Appl
45	55	46.2	998	2	US-08-702-367A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-673-789-4
Sequence 4, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
; APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
; APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
; APPLICANT: GEORGE, F.
; TITLE OF INVENTION: BSK RECEPTOR LIKE
; TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,789
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/177,812
; FILING DATE: 04-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 982
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
US-08-673-789-4

Query Match 100.0%; Score 119; DB 2; Length 982;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1   VNLLDSKTIQGELGWISYPSHG 22
            ||||||||||||||||||||||
Db   29   VNLLDSKTIQGELGWISYPSHG 50

RESULT 2
US-08-162-809-16
; Sequence 16, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; APPLICANT: Sajjadi, Fereydoun G.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,809
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-162-809-16

Qy   1   VNLLDSKTIQGELGWISYPSHG 22
            ||||||||||||||||||||||
Db   29   VNLLDSKTIQGELGWISYPSHG 50

Query Match   100.0%; Score 119; DB 1; Length 983;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches   22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
US-08-167-919A-10
; Sequence 10, Application US/08167919A
; Patent No. 5674691
; GENERAL INFORMATION:
; APPLICANT: Boyd, Andrew W.
; APPLICANT: Simpson, Richard J.
; APPLICANT: Wicks, Ian
; APPLICANT: Ward, Larry D.
; APPLICANT: Wilkinson, David
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza

; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,919A
; FILING DATE: 18-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK6841 (AU)
; FILING DATE: 21-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK9992 (AU)
; FILING DATE: 12-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00294
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9159
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-167-919A-10

Qy   1   VNLLDSKTIQGELGWISYPSHG 22
            ||||||||||||||||||||||
Db   30   VNLLDSKTIQGELGWISYPSHG 51

Query Match   100.0%; Score 119; DB 1; Length 983;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches   22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
US-08-449-645A-21
; Sequence 21, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
    NAME: Winter, Robert B.
    REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 983 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: protein
US-449-645A-21

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
       |||||||||||||||||||||||
Db 30 VNLLDSKTIQGELGWISYPSHG 51

Query Match   100.0%;   Score 119;   DB 2;   Length 983;
Best Local Similarity   100.0%;   Pred. No. 1.7e-10;
Matches   22;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 5
US-08-702-367A-21
Sequence 21, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
        Kinases
    NUMBER OF SEQUENCES: 43
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Amgen Patent Operations/RBW
        STREET: 1840 Dehaviland Drive
        CITY: Thousand Oaks
        STATE: California
        COUNTRY: USA
        ZIP: 91320
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/702,367A
        FILING DATE:
    CLASSIFICATION: 435
    ATTORNEY/AGENT INFORMATION:
        NAME: Winter, Robert B.
        REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 983 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-702-367A-21

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
       |||||||||||||||||||||||
Db 30 VNLLDSKTIQGELGWISYPSHG 51

Query Match   100.0%;   Score 119;   DB 2;   Length 983;
Best Local Similarity   100.0%;   Pred. No. 1.7e-10;
Matches   22;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 6
US-08-715-106-10
Sequence 10, Application US/08715106
Patent No. 6020306

GENERAL INFORMATION:
    APPLICANT: Boyd, Andrew W.
    APPLICANT: Simpson, Richard J.
    APPLICANT: Wicks, Ian
    APPLICANT: Ward, Larry D.
    APPLICANT: Wilkinson, David
    TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
        AND USE THEREOF
    NUMBER OF SEQUENCES: 14
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Scully, Scott, Murphy & Presser
        STREET: 400 Garden City Plaza
        CITY: Garden City
        STATE: New York
        COUNTRY: U.S.A.
        ZIP: 11530
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/715,106
        FILING DATE: 18-SEP-1996
    CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/167,919
        FILING DATE: 18-APR-1994
        APPLICATION NUMBER: PK6841 (AU)
        FILING DATE: 21-JUN-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PK9992 (AU)
        FILING DATE: 12-DEC-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PCT/AU92/00294
        FILING DATE: 19-JUN-1992
    ATTORNEY/AGENT INFORMATION:
        NAME: DiGiglio, Frank S.
        REGISTRATION NUMBER: 31,346
        REFERENCE/DOCKET NUMBER: 9159
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 742-4343
        TELEFAX: (516) 742-4366
        TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 10:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 983 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-715-106-10

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
       |||||||||||||||||||||||
Db 30 VNLLDSKTIQGELGWISYPSHG 51

Query Match   100.0%;   Score 119;   DB 3;   Length 983;
Best Local Similarity   100.0%;   Pred. No. 1.7e-10;
Matches   22;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 7
PCT-US95-04681-21
Sequence 21, Application PC/TUS9504681
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
        Kinases
    NUMBER OF SEQUENCES: 28
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Amgen Patent Operations/RBW
        STREET: 1840 Dehaviland Drive

CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||||||||||||||||||||
Db   30 VNLLDSKTIQGELGWISYPSHG 51
```

Query Match   100.0%;   Score 119;   DB 4;   Length 983;
Best Local Similarity   100.0%;   Pred. No. 1.7e-10;
Matches   22;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

PCT-US95-04681-21

RESULT 8
US-08-673-789-2
Sequence 2, Application US/0867378 9
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,789
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 877
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-2

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        |||||||:|: |:||||||:| |
Db   63 VNLLDSRTVMGDLGWIAFPKNG 84
```

Query Match   74.8%;   Score 89;   DB 2;   Length 877;
Best Local Similarity   63.6%;   Pred. No. 6.8e-06;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

RESULT 9
US-08-449-645A-30
Sequence 30, Application US/0844964 5A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 967 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-449-645A-30

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||||:|: :|:||||||: ||
Db   13 VNLLDSRVMGDLGWIAFPKNG 34
```

Query Match   74.8%;   Score 89;   DB 2;   Length 967;
Best Local Similarity   63.6%;   Pred. No. 7.6e-06;
Matches   14;   Conservative   6;   Mismatches   2;   Indels   0;   Gaps   0;

RESULT 10
US-08-702-367A-30
Sequence 30, Application US/0870236 7A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.

```
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Amgen Patent Operations/RBW
    STREET: 1840 Dehavilland Drive
    CITY: Thousand Oaks
    STATE: California
    COUNTRY: USA
    ZIP: 91320
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/702,367A
    FILING DATE:
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: Winter, Robert B.
    REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 30:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 967 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-702-367A-30

Query Match    74.8%;   Score 89;  DB 2;  Length 967;
Best Local Similarity  63.6%;  Pred. No. 7.6e-06;
Matches 14;  Conservative 6;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1  VNLLDSKTIQGELGWISYPSHG 22
         |||||||:|:|:||||:|:| :|
Db   13  VNLLDSRTVMGDLGWIAFPKNG 34


RESULT 11
US-08-449-645A-13
Sequence 13, Application US/0844964S A
Patent No. 5981245
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
    TITLE OF INVENTION: Kinases
    NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Amgen Patent Operations/RBW
    STREET: 1840 Dehavilland Drive
    CITY: Thousand Oaks
    STATE: California
    COUNTRY: USA
    ZIP: 91320
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/449,645A
    FILING DATE:
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: Winter, Robert B.
    REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 13:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 991 amino acids
    TYPE: amino acid
```

```
    TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-449-645A-13

Query Match    74.8%;   Score 89;  DB 2;  Length 991;
Best Local Similarity  63.6%;  Pred. No. 7.8e-06;
Matches 14;  Conservative 6;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1  VNLLDSKTIQGELGWISYPSHG 22
         |||||||:|:|:||||:|:| :|
Db   37  VNLLDSRTVMGDLGWIAFPKNG 58


RESULT 12
US-08-702-367A-13
Sequence 13, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
    TITLE OF INVENTION: Kinases
    NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Amgen Patent Operations/RBW
    STREET: 1840 Dehavilland Drive
    CITY: Thousand Oaks
    STATE: California
    COUNTRY: USA
    ZIP: 91320
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/702,367A
    FILING DATE:
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: Winter, Robert B.
    REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 13:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 991 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-702-367A-13

Query Match    74.8%;   Score 89;  DB 2;  Length 991;
Best Local Similarity  63.6%;  Pred. No. 7.8e-06;
Matches 14;  Conservative 6;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1  VNLLDSKTIQGELGWISYPSHG 22
         |||||||:|:|:||||:|:| :|
Db   37  VNLLDSRTVMGDLGWIAFPKNG 58


RESULT 13
PCT-US95-04681-13
Sequence 13, Application PCT/US9504681
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
    TITLE OF INVENTION: Kinases
    NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Amgen Patent Operations/RBW
    STREET: 1840 Dehavilland Drive
    CITY: Thousand Oaks
    STATE: California
```

RESULT 14
US-08-442-248-2
Sequence 2, Application US/08442248
Patent No. 5759863
GENERAL INFORMATION:
　APPLICANT: Caras, Ingrid W.
　APPLICANT: Winslow, John W.
　TITLE OF INVENTION: Al-1 Neurotrophic Factor
　NUMBER OF SEQUENCES: 18
　CORRESPONDENCE ADDRESS:
　　ADDRESSEE: Genentech, Inc.
　　STREET: 460 Point San Bruno Blvd
　　CITY: South San Francisco
　　STATE: California
　　COUNTRY: USA
　　ZIP: 94080
　COMPUTER READABLE FORM:
　　MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
　　COMPUTER: IBM PC compatible
　　OPERATING SYSTEM: PC-DOS/MS-DOS
　　SOFTWARE: patin (Genentech)
　CURRENT APPLICATION DATA:
　　APPLICATION NUMBER: US/08/442,248
　　FILING DATE: 15-MAY-1995
　　CLASSIFICATION: 435
　PRIOR APPLICATION DATA:
　　APPLICATION NUMBER: 08/330128
　　FILING DATE: 27-OCT-1994
　ATTORNEY/AGENT INFORMATION:
　　NAME: Torchia, Timothy E.
　　REGISTRATION NUMBER: 36,700
　　REFERENCE/DOCKET NUMBER: 920C4
　TELECOMMUNICATION INFORMATION:
　　TELEPHONE: 415/225-8674
　　TELEFAX: 415/952-9881
　　TELEX: 910/371-7168
　INFORMATION FOR SEQ ID NO: 2:
　　SEQUENCE CHARACTERISTICS:
　　　LENGTH: 928 amino acids
　　　TYPE: amino acid
　　　TOPOLOGY: linear

Qy　　1　VNLLDSKTIQGELGWISYPSHG 22
　　　　　||||||:|:||||||:|:|
Db　37　VNLLDSRTVMGDLGWIAFPKNG 58

Query Match　　　　　74.8%;　　Score 89;　DB 4;　Length 991;
Best Local Similarity　63.6%;　Pred. No. 7.8e-06;
Matches　14;　Conservative　6;　Mismatches　2;　Indels　0;　Gaps　0;

US-08-442-248-2

Qy　　1　VNLLDSKTIQGELGWISYPSHG 22
　　　　　||||||:|:||||||:|:|
Db　63　VNLLDSRTVLGDLGWIAFPKNG 84

Query Match　　　　　73.1%;　　Score 87;　DB 1;　Length 928;
Best Local Similarity　63.6%;　Pred. No. 1.5e-05;
Matches　14;　Conservative　6;　Mismatches　2;　Indels　0;　Gaps　0;

RESULT 15
US-08-440-815-2
Sequence 2, Application US/08440815
Patent No. 5798448
GENERAL INFORMATION:
　APPLICANT: Caras, Ingrid W.
　APPLICANT: Winslow, John W.
　TITLE OF INVENTION: Al-1 Neurotrophic Factor
　NUMBER OF SEQUENCES: 18
　CORRESPONDENCE ADDRESS:
　　ADDRESSEE: Genentech, Inc.
　　STREET: 460 Point San Bruno Blvd
　　CITY: South San Francisco
　　STATE: California
　　COUNTRY: USA
　　ZIP: 94080
　COMPUTER READABLE FORM:
　　MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
　　COMPUTER: IBM PC compatible
　　OPERATING SYSTEM: PC-DOS/MS-DOS
　　SOFTWARE: patin (Genentech)
　CURRENT APPLICATION DATA:
　　APPLICATION NUMBER: US/08/440,815
　　FILING DATE: 15-MAY-1995
　　CLASSIFICATION: 435
　PRIOR APPLICATION DATA:
　　APPLICATION NUMBER: 08/330128
　　FILING DATE: 27-OCT-1994
　ATTORNEY/AGENT INFORMATION:
　　NAME: Torchia, Timothy E.
　　REGISTRATION NUMBER: 36,700
　　REFERENCE/DOCKET NUMBER: 920C1
　TELECOMMUNICATION INFORMATION:
　　TELEPHONE: 415/225-8674
　　TELEFAX: 415/952-9881
　　TELEX: 910/371-7168
　INFORMATION FOR SEQ ID NO: 2:
　　SEQUENCE CHARACTERISTICS:
　　　LENGTH: 928 amino acids
　　　TYPE: amino acid
　　　TOPOLOGY: linear

Qy　　1　VNLLDSKTIQGELGWISYPSHG 22
　　　　　||||||:|:||||||:|:|
Db　63　VNLLDSRTVLGDLGWIAFPKNG 84

Search completed: May　9, 2000, 22:25:59
Job time: 4148 sec

COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
　MEDIUM TYPE: floppy disk
　COMPUTER: IBM PC compatible
　OPERATING SYSTEM: PC-DOS/MS-DOS
　SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
　APPLICATION NUMBER: PCT/US95/04681
　FILING DATE:
　CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
　NAME: Winter, Robert B.
　REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 13:
　SEQUENCE CHARACTERISTICS:
　　LENGTH: 991 amino acids
　　TYPE: amino acid
　　TOPOLOGY: linear
　　MOLECULE TYPE: protein
PCT-US95-04681-13

Qy　　1　VNLLDSKTIQGELGWISYPSHG 22
　　　　　||||||:|:||||||:|:|
Db　37　VNLLDSRTVMGDLGWIAFPKNG 58

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May   9, 2000, 21:52:01 ; Search time 50.71 Seconds
                   (without alignments)
                   10.276 Million cell updates/sec

Title: US-09-104-340-2
Perfect score: 119
Sequence: 1 VNLLDSKTIQGELGWISYPSHG 22

Scoring table: BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	983	1	R31466	HEK polypeptide. R
2	119	100.0	983	1	R75711	Eph-related PTK Ce
3	89	74.8	877	1	W71628	Mouse Bsk receptor
4	89	74.8	991	1	R85090	EPH-like receptor
5	87	73.1	928	1	R97853	Rat REK7 eph-relat
6	87	73.1	1005	1	W83147	Rat receptor tyros
7	79	66.4	986	1	W85936	Protein tyrosine-k
8	79	66.4	986	1	R85091	EPH-like receptor
9	73	61.3	948	1	W83148	Rat receptor tyros
10	62	52.1	977	1	W12258	Embryonic stem cel
11	61	51.3	998	1	R85092	EPH-like receptor
12	59	49.6	610	1	W03422	Mouse developmenta
13	59	49.6	626	1	W03423	Mouse developmenta
14	59	49.6	998	1	W03421	Mouse developmenta
15	57	47.9	968	1	W19968	Human alanyl-tRNA
16	57	47.9	1004	1	Y07088	Renal cancer assoc
17	57	47.9	970	1	R85089	EPH-like receptor
18	56	47.1	994	1	W87018	Receptor tyrosine
19	56	47.1	994	1	W26356	Mouse Nuk tyrosine
20	56	47.1	995	1	R75712	Eph-related PTK Ce
21	56	47.1	1011	1	R75709	Eph-related PTK Ce
22	55	46.2	984	1	R44513	elk. Expression of
23	55	46.2	990	1	R51899	Human embryonal ki
24	55	46.2	951	1	R75704	Eph-related CEK6.
25	52	43.7	1006	1	W72256	Human receptor typ
26	52	43.7	1006	1	W70525	Human thymus recep
27	52	43.7	1021	1	W70526	Human thymus recep
28	48	40.3	973	1	W75707	Eph-related prk Ce
29	46	38.7	619	1	W75771	Human GTP binding
30	45	37.8	530	1	R50972	Norwalk virus stra
31	45	37.8	530	1	W32213	Mutanase of penici
32	45	37.8	630	1	W89392	Penicillium purpur
33	45	37.8	630	1	W97037	Penicillium purpur
34	43	36.1	530	1	R57091	Small round virus
35	43	36.1	573	1	W62295	Glutathionylsperm1
36	43	35.3	330	1	R98599	Broccoli ACC synth
37	42	35.3	333	1	R46299	Homologue of flavo
38	42	35.3	335	1	R46300	Homologue of flavo
39	42	35.3	374	1	R98898	Broccoli ACC synth
40	41.5	34.9	426	1	R43300	Broccoli ACC synth
41	41	34.5	122	1	R10097	Xenopus Bone Morph
41	41	34.5	122	1	W22109	Amino acid sequenc
42	41	34.5	276	1	W28109	Human NES1 polypep
43	41	34.5	483	1	W07620	Secretory signal p
44	41	34.5	958	1	R07445	Yeast alanyl-tRNA
45	40.5	34.0	511	1	W19969	Helicobacter polyp

ALIGNMENTS

RESULT 1

ID   R31466 standard; Protein; 983 AA.
AC   R31466;
DT   24-MAY-1993 (first entry)
DE   HEK polypeptide.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; JM; epithelial; B;cellular response; growth; differentiation.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FH
FT   peptide   1..20
FT   /note= "Signal peptide"
FT   modified_site   21..39
FT   /note= "N-link glycosylation"
FT   modified_site   337..339
FT   /note= "N-link glycosylation"
FT   modified_site   391..393
FT   /note= "N-link glycosylation"
FT   protein   404..406
FT   /note= "Purified HEK protein #1"
FT   modified_site   493..495
FT   /note= "N-link glycosylation"
FT   domain   542..565
FT   /note= "N-link glycosylation"
FT   binding_site   628
FT   /note= "Transmembrane domain"
FT   binding_site   630
FT   /note= "ATP binding site"
FT   binding_site   633
FT   /note= "ATP binding site"
FT   modified_site   779
FT   /note= "ATP binding site"
FT   protein   840..860
FT   /note= "Purified HEK protein #2"
FT   modified_site   232..234
FT   /note= "Putative autophosphorylation site"
PN   WO9300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU0294.
PR   19-JUN-1992; AU-006841.
PR   21-JUN-1991; AU-009992.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   WPI; 93-036373/04.
DR   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence represets human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the

CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence   983 AA;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
          ||||||||||||||||||||||
Db   30 VNLLDSKTIQGELGWISYPSHG 51

RESULT   2
ID   R75711 standard; Protein; 983 AA.
AC   R75711;
DT   11-NOV-1995   (first entry)
DE   Eph-related PTK Cek4.
KW   Cek4; Eph; protein tyrosine-kinase; cancer;
KW   prognosis.
OS   Gallus sp.
PN   WO9515375-A.
PD   08-JUN-1995.
PF   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;
DR   WPI; 95-215256/28.
DR   N-PSDB; Q90659.
CC   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
CC   cancer.
PS   Disclosure; Page 85-89; 129pp; English.
PT   Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
PT   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
PT   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and
CC   embryonic tissues and also in the adult brain and retina.
SQ   Sequence   983 AA;

Query Match   100.0%; Score 119; DB 1; Length 983;
Best Local Similarity 100.0%;   Pred. No.1.1e-10;
Matches   22; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
          ||||||||||||||||||||||
Db   29 VNLLDSKTIQGELGWISYPSHG 50

RESULT   3
ID   W71628 standard; Protein; 877 AA.
AC   W71628;
DT   25-NOV-1998   (first entry)
DE   Mouse Bsk receptor-like tyrosine kinase.
KW   Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW   neurodegenerative disease; limbic system neuron regeneration;
KW   chromosomal abnormality; degenerative growth; development disorder;
KW   viral infection; bacterial infection; Alzheimer's disease; epilepsy;
OS   Mus sp.
PN   US581479-A.
PD   29-SEP-1998.
PF   11-JUN-1996; 673789.
PR   04-JAN-1994; US-177812.
PR   11-JUN-1996; US-673789.
PA   (KROM/) KROMER L F.
PA   (SCHU/) SCHULZ N T.
PA   (WOUD/) WOUDE G F V.
PA   (ZHOU/) ZHOU R.

PI   Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR   WPI; 98-541751/46.
DR   N-PSDB; V58192.
CC   Isolated nucleic acid sequence encoding protein - used in Bsk
CC   tyrosine kinase. The present sequence represents mouse Bsk, which is a receptor-like
CC   Bsk nucleic acid probes, which can be used in detecting alterations in level of Bsk
CC   the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC   from a mammal afflicted with a disease, such as neurodegenerative
CC   diseases or disorders and neoplasms. The nucleic acid sequence can also
CC   be delivered into the limbic system of patients with limbic system
CC   neurodegenerative disease, disorder or injury, to promote or enhance
CC   limbic system neuron regeneration or growth. Such neurodegenerative
CC   diseases include, chromosomal abnormalities, degenerative growth and
CC   development disorders, viral infections, bacterial infections, brain
CC   injuries, neoplastic conditions, Alzheimer's disease, epilepsy,
CC   schizophrenia, or stroke and cerebral ischaemia.
PS   Claim 1; Fig 2; 72pp; English.
PT   The present sequence represents mouse Bsk, which is a receptor-like
PT   nucleic acid probes, used in detecting alterations in level of Bsk
PT   messenger-RNA in biological samples isolated from mammal afflicted
PT   with disease
SQ   Sequence   877 AA;

Query Match   74.8%; Score 89; DB 1; Length 877;
Best Local Similarity   63.6%;   Pred. No. 5.2e-06;
Matches   14; Conservative   6; Mismatches   2; Indels   0; Gaps   0;

Qy   1 VNLLDSKTIQGELGWISYPSHG 22
          ||||||:|:| :||||||:|:|:|
Db   63 VNLLDSRTVMGDLGWIAFPKNG 84

RESULT   4
ID   R85090 standard; Protein; 991 AA.
AC   R85090;
DT   16-APR-1996   (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK7.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK7;
KW   human eph-like kinase; therapy; diagnosis; vector; antibody.
OS   Homo sapiens
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-37379/48.
DR   N-PSDB; TO2947.
CC   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
CC   and related vectors, host cells, proteins, antibodies etc; used
CC   diagnostically and therapeutically to modulate receptor activation
CC   or prodn.
PS   Claim 18; Page 54-57; 133pp; English.
PT   4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
PT   HEK8 and HEK11 (R85089-92), respectively, were identified following
PT   isolation of their encoding cDNAs (TO2946-49) from a human foetal brain
PT   cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC   HEK receptors (or their soluble extracellular domains) are produced by
CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   assays), or to identify ligands for HEK receptors, or
CC   therapeutically to modulate the activation of cell-associated
CC   receptors. Soluble HEK7 receptor may primarily affect
CC   proliferation and/or differentiation of brain cells.
CC   pancreatic cells.
SQ   Sequence   991 AA;

Query Match   74.8%; Score 89; DB 1; Length 991;

RESULT 5

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||::|:|:||||:||:|  :|
Db   37 VNLLDSRTVMGDLGWIAFPKNG 58
```

Best Local Similarity 63.6%; Pred. No. 6e-06;
Matches 14; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

```
ID   R97853 standard; Protein; 928 AA.
AC   R97853;
DT   05-JAN-1997 (first entry)
DE   Rat REK7 eph-related tyrosine kinase receptor.
KW   REK7; eph-related tyrosine kinase receptor; Al-1; ligand;
KW   neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis.
OS   Rattus sp.
FH   
FT   Key            Location/Qualifiers
FT   peptide        1..57
FT                  /label= Sig_peptide
FT   protein        58..928
FT                  /label= Mat_protein
FT   domain         58..462
FT                  /label= Extracellular_domain
PN   WO9613518-A1.
PD   09-MAY-1996.
PF   26-OCT-1995; U14016.
PR   27-OCT-1994; US-330128.
PR   07-JUN-1995; US-486449.
PA   (GETH ) GENENTECH INC.
PI   Caras IW, Winslow JW;
WPI; 96-239448/24.
DR   N-PSDB; T18893.
CC   AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
CC   treatment and diagnosis of neuronal disorders and
PS   angiogenesis-related conditions.
PT   Example 1; Page 50-53; 75pp; English.
PT   Rat REK7 (R97853) is an eph-related tyrosine kinase receptor, for
CC   which AL-1 (see also W97854) is a ligand. Its amino acid sequence
CC   was deduced from a cDNA clone (T18893) isolated from a hippocampal
CC   cDNA library. An REK-IgG fusion was used to screen cultured cell
CC   lines for surface expression of REK7-binding activity. Primers
CC   (see also T44382-83) based on peptide sequences (R97856-59) of
CC   isolated ligands were used to amplify human breast carcinoma BT20
CC   cell cDNA, and an amplified fragment was used to screen a human
CC   foetal brain cDNA library, leading to the isolation of AL-1 cDNA
CC   (T18897).
SQ   Sequence   928 AA;
```

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||::|:|:||||:||:|  :|
Db   63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Query Match 73.1%; Score 87; DB 1; Length 928;
Best Local Similarity 63.6%; Pred. No. 1.1e-05;
Matches 14; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

RESULT 6

```
ID   W83147 standard; Protein; 1005 AA.
AC   W83147;
DT   11-FEB-1999 (first entry)
DE   Rat receptor tyrosine kinase Ehk-1.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis.
OS   Rattus sp.
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995; 469537.
PR   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Masiakowski P, Yancopoulos GD;
WPI; 99-044584/04.
N-PSDB; V70207.
DR   
PT   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Example; Fig 22; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence represents rat Ehk-1.
SQ   Sequence  1005 AA;
```

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||::|:|:||||:||:|  :|
Db   63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Query Match 73.1%; Score 87; DB 1; Length 1005;
Best Local Similarity 63.6%; Pred. No. 1.3e-05;
Matches 14; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

RESULT 7

```
ID   R85936 standard; Protein; 986 AA.
AC   R85936;
DT   14-FEB-1996 (first entry)
DE   Protein tyrosine kinase bpTK7.
KW   Protein tyrosine-kinase; pTK; bpTK7; agonist; cell growth;
KW   differentiation.
OS   Homo sapiens.
FH   
FT   Key            Location/Qualifiers
FT   peptide        1..19
FT                  /label= Sig_peptide
FT   domain         20..547
FT                  /label= Extracellular_domain
FT   domain         548..570
FT                  /label= Transmembrane_domain
FT   domain         571..986
FT                  /label= Intracellular_tyrosine-kinase_domain
PN   WO9527061-A1.
PD   12-OCT-1995.
PF   04-APR-1995; U04228.
PR   04-APR-1994; US-222616.
PA   (GETH ) GENENTECH INC.
PI   Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
Wood WI;
WPI; 95-361160/47.
N-PSDB; T03100.
DR   
PT   Agonist antibodies which activate specific protein tyrosine
PT   kinase(s) - also activate chimeric proteins of kinase extracellular
PT   domain and Ig constant domain, useful for studying, and therapeutic
PT   modulation of, cell growth and differentiation
PS   Disclosure; Page 95-99; 125pp; English.
CC   DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC   to screen cDNA libraries to identify novel pTK genes. The bpTKs,
CC   bpTK1, bpTK2, bpTK3, bpTK4, bpTK5 and bpTK7 (R85924-28 and R85935,
CC   respectively) are expressed in human brain tissue and show homology
CC   to known pTKs. A full-length sequence for bpTK7 (R85936) was also
CC   obtd. bpTK7 can be used to design drugs that modulate pTK activity.
SQ   Sequence   986 AA;
```

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        ||||::|::|:||||||||:| |
Db   63 VNLLDSRTVLGDLGWIAFPKNG 84
```

Query Match 66.4%; Score 79; DB 1; Length 986;
Best Local Similarity 63.6%; Pred. No. 0.00023;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy    1 VNLLDSKTIQGELGWISYPSHG 22
        |||||::|::|:|||||||:| |
```

```
Db    31 VTLLDSRSVQGELGWIASPLEG 52

RESULT   8

ID    R85091 standard; Protein; 986 AA.
AC    R85091;
DT    16-APR-1996 (first entry)
DE    EPH-like receptor protein tyrosine kinase HEK8;
DE    EPH-like receptor protein tyrosine kinase HEK8;
KW    EPH-like receptor protein tyrosine kinase; PTK; HEK8;
KW    human eph-like kinase; therapy; diagnosis; vector; antibody.
OS    Homo sapiens.
PN    WO9528484-A1.
PD    26-OCT-1995.
PF    14-APR-1995; U04681.
PR    15-APR-1994; U04681.
PR    15-APR-1994; US-229509.
PA    (AMGE-) AMGEN INC.
PI    Fox GM, Jing S, Welcher AA;
DR    WPI; 95-373799/48.
DR    N-PSDB; T02948.
PT    New nucleic acid encoding EPH-like receptor protein tyrosine kinase(s) -
PT    and related vectors, host cells, proteins, antibodies etc., used
PT    diagnostically and therapeutically to modulate receptor activation
PT    or prodn.
PS    Claim 18; Page 62-65; 133pp; English.
CC    4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC    HEK8 and HEK11 (R85089-92), respectively, were identified following
CC    isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC    cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC    catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC    HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC    HEK receptors (or their soluble extracellular domains) are produced by
CC    expression of encoding sequences in procaryotic or eucaryotic host
CC    cells, and are used to produce antibodies (utilised in diagnostic
CC    assays), or to identify and purify ligands for HEK receptors, or
CC    therapeutically to modulate the activation of cell-associated
CC    receptors.
SQ    Sequence   986 AA;

Qy    Query Match     66.4%;   Score 79;  DB 1;  Length 986;
      Best Local Similarity   63.6%;   Pred. No. 0.00023;
      Matches   14;  Conservative   4;  Mismatches   4;  Indels   0;  Gaps   0;

Qy       1  VNLLDSKTIQGELGWISYPSHG  22
            ||||:::||||||||||:|:|
Db      31  VTLLDSRSVQGELGWIASPLEG  52

RESULT   9

ID    W83148 standard; Protein; 948 AA.
AC    W83148;
DT    11-FEB-1999 (first entry)
DE    Rat receptor tyrosine kinase Ehk-2.
DE    Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW    neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW    binding protein; BDNF; NT-3; diagnosis.
OS    Rattus sp.
PN    US5843749-A.
PD    01-DEC-1998.
PF    06-JUN-1995; 469537.
PR    06-JUN-1995; US-406247.
PR    17-MAR-1995; US-406247.
PR    26-JUL-1991; US-736559.
PR    28-OCT-1993; US-144992.
PR    06-JUN-1995; US-469537.
PA    (REGE-) REGENERON PHARM INC.
PI    Maisonpierre PC, Maslakowski P, Yancopoulos GD;
DR    WPI; 99-044584/04.
DR    N-PSDB; V70208.
PT    DNA encoding receptor tyrosine kinase proteins - and corresponding
PT    proteins
PS    Example: Fig 21; 194pp; English.
```

```
CC    The present invention describes nucleic acid molecules for ror-1,
CC    ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC    Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC    tyrosine kinases. The present sequence represents rat Ehk-2.
SQ    Sequence   948 AA;

Qy    Query Match     61.3%;   Score 73;  DB 1;  Length 948;
      Best Local Similarity   59.1%;   Pred. No. 0.0019;
      Matches   13;  Conservative   4;  Mismatches   5;  Indels   0;  Gaps   0;

Qy       1  VNLLDSKTIQGELGWISYPSHG  22
            |||||::|||||||||:|:|
Db      34  VVLLDTSTVMGELGWKIYPLHG  55

RESULT  10

ID    W19258 standard; Protein; 977 AA.
AC    W19258;
DT    15-SEP-1997 (first entry)
DE    Embryonic stem cell kinase (Esk).
DE    Embryonic stem cell kinase; receptor tyrosine kinase; mucositis;
KW    epithelium; signal transduction; gene therapy; diagnosis.
OS    Mus musculus.

FH    Key              Location/Qualifiers
FT    peptide          1..26
FT                     /note= "Sig_peptide"
FT    protein          27..977
FT                     /label= Mat_protein
FT    protein          27..548
FT                     /label= Extracellular_domain
FT    domain           549..569
FT                     /label= Transmembrane_domain
FT    domain           570..977
FT                     /label= Intracellular_domain
FT    misc_difference  70
FT                     /note= "conserved Cys residue"
FT    misc_difference  107
FT                     /note= "conserved Cys residue"
FT    peptide          119
FT                     /note= "conserved Cys residue"
FT    misc_difference  192
FT                     /note= "conserved Cys residue"
FT    protein          205
FT                     /note= "conserved Cys residue"
FT    domain           232
FT                     /note= "conserved Cys residue"
FT    domain           250
FT                     /note= "conserved Cys residue"
FT    misc_difference  263
FT                     /note= "conserved Cys residue"
FT    misc_difference  265
FT                     /note= "conserved Cys residue"
FT    misc_difference  278
FT                     /note= "conserved Cys residue"
FT    misc_difference  281
FT                     /note= "conserved Cys residue"
FT    misc_difference  295
FT                     /note= "conserved Cys residue"
FT    misc_difference  298
FT                     /note= "conserved Cys residue"
FT    misc_difference  312
FT                     /note= "conserved Cys residue"
FT    misc_difference  314
FT                     /note= "conserved Cys residue"
FT    misc_difference  330
FT                     /note= "conserved Cys residue"
FT    misc_difference  330
FT                     /note= "conserved Cys residue"
FT    misc_difference  384
FT                     /note= "conserved Cys residue"
FT    misc_difference  387
FT                     /note= "conserved Cys residue"
FT    peptide          331..441
```

RESULT 11
ID   R85092
AC   R85092;
DT   16-APR-1996 (first entry)
Db   29 VTLMDTSTAQGELGWLLDP 47
Qy   1 VNLLDSKTIQGELGWISYP 19
           |:|:| : | ||||||:| |
Matches   11; Conservative   3; Mismatches   5; Indels   0; Gaps   0;
Best Local Similarity   57.9%; Pred. No. 0.11;
Query Match   52.1%;   Score 62; DB 1; Length 977;

SQ   Sequence   977 AA;
CC   infections and pancreatitis.
CC   Sjorgren's syndrome and related autoimmune disease, mucosal
CC   corneal or peptic ulcers, ulcerative colitis, oesophageal reflux,
CC   modulation may be used to treat e.g. skin injuries, hair loss,
CC   treatment), and as Esk is expressed in all epithelial tissues its
CC   can be used to treat mucositis (partic. when caused by cancer
CC   reagents or therapeutic antagonists. Modulation of Esk expression
CC   as immunogens for prodn. of antibodies useful as diagnostic
CC   for ligand isolation, can be produced in host cells and used
CC   Esk (extracellular domain), can be produced in host cells and used
CC   liver cDNA library. Recombinant Esk polypeptides, esp. soluble
CC   was deduced from a DNA molecule (T72320) isolated from a murine
CC   kinase that belongs to the Eph subfamily. Its amino acid sequence
CC   Novel mouse embryonic stem cell kinase (Esk) is a receptor tyrosine
Claim 1; Page 48-54; 98pp; English.
PS   Claim 1; Page 48-54; 98pp; English.
PT   mucositis and other disorders involving epithelium
PT   for production of modulators and antibodies, useful to treat
PT   DNA encoding embryonic stem cell kinase, receptor tyrosine
N-PSDB; T72320.
DR   N-PSDB; T72320.
WPI; 97-351065/32.
PI   Boyd AW, Lickliter J;
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PR   22-DEC-1995; AU-007277.
PD   20-DEC-1996; AU0826.
PN   WO9723629-A1.

FT   /note= "potential N-linked glycosylation site"
FT   /label- Glycosylation
FT   479
FT modified_site
FT   /note= "potential N-linked glycosylation site"
FT   /label- Glycosylation
FT   415
FT modified_site
FT   /note= "potential N-linked glycosylation site"
FT   /label- Glycosylation
FT   339
FT modified_site
FT   /note= "potential N-linked glycosylation site"
FT   /label- Glycosylation
FT   60
FT modified_site
FT   /note= "PIRWTAPE motif indicating substrate
FT   specificity for tyrosine"
FT   790..797
FT peptide
FT   /note= "DLAARN motif indicating substrate
FT   specificity for tyrosine"
FT   750..755
FT peptide
FT   /note= "invariant lysine residue associated with
FT   ATP binding motif"
FT   657
FT binding_site
FT   /note= "GXGXXG ATP binding motif"
FT   632..637
FT binding_site
FT   /note= "characteristic of Eph subfamily"
FT   448..536
FT peptide
FT   /note= "Fibronectin_type_III_repeat"
FT   /label- Fibronectin_type_III_repeat

RESULT 12
ID   W03422 standard; Protein; 610 AA.
AC   W03422;
DT   11-NOV-1996 (first entry)
Db   33 VLLLDSKAQQTELEWISSPPNG 54
Qy   1 VNLLDSKTIQGELGWISYPSHG 22
           |||||||| ||||:|| | :|
Matches   14; Conservative   1; Mismatches   7; Indels   0; Gaps   0;
Best Local Similarity   63.6%; Pred. No. 0.16;
Query Match   51.3%;   Score 61; DB 1; Length 998;

SQ   Sequence   998 AA;
CC   receptors.
CC   therapeutically to modulate the activation of cell-associated
CC   assays), or to identify and purify ligands for HEK receptors, or
CC   cells, and are used to produce antibodies (utilised in diagnostic
CC   expression of encoding sequences in procaryotic or eucaryotic host
CC   HEK receptors (or their soluble extracellular domains) are produced by
CC   HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC   catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC   cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC   isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC   HEK8 and HEK11 (R85089-92), respectively, were identified following
CC   4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
PS   Claim 18; Page 71-75; 133pp; English.
PT   or prodn.
PT   diagnostically and therapeutically to modulate receptor activation
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   New nucleic acid encoding EPH-like receptor protein tyrosine kinase(s)
N-PSDB; T02949.
DR   N-PSDB; T02949.
WPI; 95-373799/48.
PI   Fox GM, Jing S, Welcher AA;
PA   (AMGE-) AMGEN INC.
PR   15-APR-1994; US-229509.
PD   14-APR-1995; U04681.
PN   WO9528484-A1.
OS   Homo sapiens.
KW   human eph-like kinase; therapy; diagnosis; antibody; vector.
DE   EPH-like receptor protein tyrosine kinase; PTK; HEK11;
DE   EPH-like receptor protein tyrosine kinase HEK11.

RESULT 13
ID   W03424 standard; Protein; 610 AA.
AC   W03422;
DT   11-NOV-1996 (first entry)
Db   33 VLLLDSKTIQGELGWISYPSHG 22
Qy   1 VNLLDSKTIQGELGWISYPSHG 22
           |||||||| ||||:|| | :|
Matches   14; Conservative   1; Mismatches   7; Indels   0; Gaps   0;
Best Local Similarity   63.6%; Pred. No. 0.16;
Query Match   51.3%;   Score 61; DB 1; Length 998;

SQ   Sequence   998 AA;
FT   /note= "product of alternative splicing"
FT   600..610
FT misc_difference
FT   /label- Transmembrane_domain
FT   555..579
FT domain
FT   /label- N-glycosylation_site
FT   410..412
FT modified_site
FT   /label- N-glycosylation_site
FT   343..345
FT modified_site
FT   /label- Sig_peptide
FT   64..66
FT modified_site
FT   /label- Sig_peptide
FT   1..28
FT peptide
FH   Key   Location/Qualifiers
OS   Mus sp.
KW   neurodegeneration; neuroproliferation; cancer.
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
DE   Mouse developmental kinase 1; MDK1 T1; receptor tyrosine kinase;
DE   Mouse developmental kinase 1 MDK1 T1.
PS   Claim 18; Page 71-75; 133pp; English.
PR   03-JAN-1995; US-368776.
PD   11-JUL-1996; U00419.
PN   WO9621013-A1.
DR   N-PSDB; T32961.
WPI; 96-333988/33.
PI   Ciossek T, Millauer B, Ullrich A;
PA   (SUGE-) SUGEN INC.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

RESULT 13
W03423
ID   W03423 standard; Protein; 626 AA.
AC   W03423;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1 MDK1 T2.
KW   Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer.
OS   Mus sp.
FH   Key   Location/Qualifiers
FH
FT   peptide   1..28
FT   /label= Sig_peptide
FT   modified_site   64..66
FT   /label= N-glycosylation_site
FT   modified_site   343..345
FT   /label= N-glycosylation_site
FT   modified_site   410..412
FT   /label= N-glycosylation_site
FT   domain   555..579
FT   /label= N-glycosylation_site
FT   misc_difference 600..626
FT   /label= Transmembrane_domain
FT   /note= "product of alternative splicing"
PN   WO9621013-A1.
PD   11-JUL-1996;
PR   03-JAN-1995; US-368776.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
WPI 96-333988/33.
DR   N-PSDB: T32962.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PS   pathways
CC   Example 1; Page 113-115; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 truncated versions, MDK1 T1 (W03422) and MDK1 T2
CC   (W03423), of the novel mouse developmental kinase 1 (MDK1) (see also
CC   W03421), a new member of the eck/eph family of receptor tyrosine
CC   kinases. Their amino acid sequences were deduced from cDNA clones
CC   (T32961 and T32962) obtd. from adult mouse brains. MDK T1 and T2
CC   each possess the entire ectodomain, the transmembrane domain and
CC   part of the juxtamembrane region of MDK1, but lack the catalytic
CC   tyrosine kinase domain. They can be used to screen for potential
CC   agents useful for treatment of diseases characterised by abnormal
CC   signal transduction.
SQ   Sequence 610 AA;

Query Match   49.6%;   Score 59;   DB 1;   Length 610;
Best Local Similarity 63.6%;   Pred. No. 0.19;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VNLLDSKTIQGELGWISYPSHG 22
       ||||| | ||||||| |
Db 33 VLLLDSKAQQTELEWISSPPSG 54

RESULT 14
W03421
ID   W03421 standard; Protein; 998 AA.
AC   W03421;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1.
KW   Mouse developmental kinase 1; MDK1; receptor tyrosine kinase; RTK;
KW   signal transduction; probe; diagnosis; therapy;
KW   neurodegeneration; neuroproliferation; cancer.
OS   Mus sp.
FH   Key   Location/Qualifiers
FH
FT   peptide   1..28
FT   /label= Sig_peptide
FT   modified_site   64..66
FT   /label= N-glycosylation_site
FT   modified_site   343..345
FT   /label= N-glycosylation_site
FT   modified_site   410..412
FT   /label= N-glycosylation_site
FT   domain   555..579
FT   /label= N-glycosylation_site
FT   /label= Transmembrane_domain
PN   WO9621013-A1.
PD   11-JUL-1996;
PR   03-JAN-1995; U00419.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
WPI 96-333988/33.
DR   N-PSDB: T32960.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
CC   Example 1; Page 105-108; 128pp; English.
CC   Mouse developmental kinase 1 (MDK1) (W03421) is a new member of the
CC   eck/eph family of receptor tyrosine kinases (RTKs). Its amino
CC   acid sequence was deduced from a cDNA clone (T32960) isolated
CC   from mouse embryo and adult brain libraries. The distinct
CC   patterns of MDK1 expression during mouse development suggest an
CC   important role for MDK1 in the formation of neuronal structures.
CC   MDK1 may be obtd. by expression in host cells. It can be used
CC   in methods for the diagnosis of diseases characterised by
CC   abnormality in a signal transduction pathway, such as
CC   neuroproliferative or neurodegenerative disorders or cancer,
CC   to screen for (ant)agonists, and to raise antibodies.
SQ   Sequence 998 AA;

Query Match   49.6%;   Score 59;   DB 1;   Length 998;
Best Local Similarity 63.6%;   Pred. No. 0.33;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VNLLDSKTIQGELGWISYPSHG 22
       ||||| | ||||||| |
Db 33 VLLLDSKAQQTELEWISSPPSG 54

RESULT 15
W19968
ID   W19968 standard; Protein; 968 AA.

```
AC   W19968;
DT   19-AUG-1997  (first entry).
DE   Human alanyl-tRNA synthetase.
KW   Alanyl-tRNA synthetase; ligase; analysis; diagnosis; autoimmune;
KW   disease; antisynthetase syndrome; interstitial lung disease;
KW   arthritis; Raynaud's phenomenon; fever; detection.
OS   Homo sapiens.
PN   US5629188-A.
PD   13-MAY-1997.
PF   21-APR-1995; 426236.
PR   21-APR-1995; US-426236.
PA   (CANC-) CANCER INST JAPANESE FOUND CANCER.
PA   (CUBI-) CUBIST PHARM INC.
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
PI   Ripmaster TL, Schimmel PR, Shiba K;
PA   WPI; 97-280271/25.
DR   N-PSDB; T71086.
PT   Human and yeast alanyl-tRNA synthetase genes - useful for production
PT   of recombinant enzymes for diagnostic and analytical applications
PS   Claim 3; Column 43-48; 42pp; English.
CC   W19968 shows the amino acid sequence of human alanyl-tRNA synthetase
CC   (Ala-tRS). Human Ala-tRS and antibodies specific for it, may be useful
CC   in assays for diagnosing and monitoring an autoimmune disease known as
CC   antisynthetase syndrome. A condition in which patients having
CC   polymyositis and/or dermatomyositis produce antibodies against the
CC   essential synthetase enzyme. Anti-synthetase syndrome is characterised
CC   by a number of clinical presentations, including interstitial lung
CC   disease, arthritis, Raynaud's phenomenon and fever. Human Ala-tRS can
CC   be used to assess the toxic effects of antimicrobial Ala-tRS inhibitors
CC   on human Ala-tRS. Tester strains comprising host cells containing
CC   recombinant Ala-tRS gene which complements or substitutes the
CC   function of the native gene can be used to test for any toxic effects
CC   of such antimicrobial agents that specifically interacts with the human
CC   Ala-tRS gene or protein. Ala-tRS proteins can be used to separate
CC   alanine from other amino acids, to separate L-alanine from D-alanine,
CC   to isolate tRNA that specifically recognises Ala-tRS, and for
CC   quantitative determination of alanine or ATP.
SQ   Sequence 968 AA;

Query Match             47.9%;   Score 57;  DB 1;  Length 968;
Best Local Similarity   47.6%;   Pred. No. 0.56;
Matches   10;  Conservative   5;  Mismatches   6;  Indels   0;  Gaps   0;

Qy        2  NLLDSKTIQGELGWISYPSHG  22
             :|||||| |:| :||| :||
Db      399  SLGDSKTIPGDTAWLLYDTYG  419
```

Search completed: May 9, 2000, 21:52:03
Job time: 3755 sec

This Page Blank (uspto)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May   9, 2000, 22:31:10 ; Search time 71.69 Seconds
                                              (without alignments)
                                              28.047 Million cell updates/sec

Title: US-09-104-340-3
Perfect score: 147
Sequence: 1 MDCQLSILLLSCSVLDSFGELIPQPSNE 29

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5
Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   SPTREMBL_12:*
                  1: sp_archea:*
                  2: sp_bacteria:*
                  3: sp_fungi:*
                  4: sp_human:*
                  5: sp_invertebrate:*
                  6: sp_mammal:*
                  7: sp_mhc:*
                  8: sp_organelle:*
                  9: sp_phage:*
                  10: sp_plant:*
                  11: sp_rodent:*
                  12: sp_virus:*
                  13: sp_vertebrate:*
                  14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	58.5	39.8	495	5	Q17572	Q17572 caenorhabdi
2	50	34.0	1022	5	O76154	O76154 dugesia jap
3	49.5	33.7	888	10	O81224	O81224 arabidopsis
4	49.5	33.7	888	10	O81225	O81225 arabidopsis
5	49.5	33.7	901	10	O80462	O80462 arabidopsi
6	49	33.3	147	11	Q23683	Q23683 caenorhabdi
7	47	32.0	349	11	Q9WTX3	Q9WTX3 rattus norv
8	47	32.0	399	5	O02232	O02232 caenorhabdi
9	47	32.0	429	11	O27520	O27520 methanobact
10	47	32.0	514	10	Q9VS8	Q9VS8 arabidopsis
11	47	32.0	1146	5	Q90Y9	Q90Y9 drosophila
12	47	32.0	1622	11	Q9Z330	Q9Z330 rattus ratt
13	46	31.3	431	3	O94559	O94559 schizosacch
14	46	31.3	793	5	O18215	O18215 caenorhabdi
15	46	31.3	2304	12	O88893	O88893 tobacco rin
16	45.5	31.0	528	10	O81339	O81339 mesembryant
17	45.5	31.0	975	5	Q19445	Q19445 caenorhabdi
18	45	30.6	79	2	P73491	P73491 synechocyst
19	45	30.6	223	5	P91851	P91851 caenorhabdi
20	45	30.6	263	12	O12297	O12297 feline coro
21	45	30.6	343	2	Q9X958	Q9X958 streptomyce
22	45	30.6	370	6	P79333	P79333 oryctolagus
23	45	30.6	563	1	Q58373	Q58373 methanococc
24	45	30.6	855	12	Q66960	Q66960 feline immu
25	45	30.6	1490	13	Q66960	Q66960 feline immu
26	45	30.6	3898	12	P79922	P79922 xenopus lae
27	45	30.6	3898	12	Q68534	Q68534 hog cholera
28	45	30.6	3898	12	Q96891	Q96891 hog cholera
29 44.5	44.5	30.3	276	4	Q13816	Q13816 homo sapien
30	44.5	30.3	111	2	Q9Z918	Q9Z918 chlamydia p
31	44	29.9	157	11	P62338	P62338 mus musculu
32	44	29.9	170	12	Q66939	Q66939 feline immu
33	44	29.9	191	12	P87266	P87266 saccharomyc
34	44	29.9	224	12	Q85538	Q85538 feline immu
35	44	29.9	286	5	O17231	O17231 caenorhabdi
36	44	29.9	302	2	Q9XBS0	Q9XBS0 zymomonas m
37	44	29.9	376	2	Q9XBS0	Q9XBS0 zymomonas m
38	44	29.9	380	4	Q92FE5	Q92FE5 vibrio para
39	44	29.9	411	10	O15885	O15885 homo sapien
40	44	29.9	482	2	Q9QR4	Q9QR4 arabidopsis
41	44	29.9	700	10	P96418	P96418 arabidopsis
42	44	29.9	734	5	O81887	O81887 mycobacteri
43	44	29.9	700	13	Q9YHD9	Q9YHD9 brachydanio
44	44	29.9	1537	13	O17237	O17237 caenorhabdi
45	44	29.9	4123	4	O75851	O75851 homo sapien

ALIGNMENTS

RESULT   1
Q17572

ID   Q17572   PRELIMINARY;   PRT;   495 AA.
AC   Q17572;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   C01G6.3 PROTEIN.
GN   C01G6.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BERKS M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z35595; CAA84634.1; -.
SQ   SEQUENCE   495 AA;   57697 MW; 5B449B81 CRC32;

Query Match   58.5; Score 58.5; DB 5; Length 495;
Best Local Similarity   39.8%; Pred. No. 0.58;
Matches   16; Conservative   6; Mismatches   3; Indels 19; Gaps   2;

Qy   1 MDCQLSILLLL-------SCSVLDSF-----------------GELIPQ 25
          :|| :| :|:|:|| :|||||||| |||:
Db   105 IDFQMRLILLLLICSAFSTCSVLDSFRANGIEFEVYGECRLIPE 148

RESULT 2

```
ID   Q76154
AC   Q76154;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   NA+/K+-ATPASE ALPHA-SUBUNIT.
GN   DJNAK.
OS   Dugesia japonica (Planarian).
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Tricladida;
OC   Paludicola; Dugesiidae; Dugesia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GI; TISSUE=WHOLE ANIMAL;
RA   OGATA S., WATANABE K.;
RT   "Planarian Na+/K+-ATPase alpha-subunit.";
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB012391; BAA32798.1; -.
DR   PFAM; PF00122; E1-E2_ATPase; 1.
DR   PFAM; PF00689; Na_K_ATPase_C; 1.
DR   PFAM; PF00690; Na_K_ATPase_N; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PRINTS; PR00121; NAKATPASE.
SQ   SEQUENCE 1022 AA; 113462 MW; 88S294F5 CRC32;

Query Match    34.0%;  Score 50;  DB 5;  Length 1022;
Best Local Similarity   41.2%;
Matches   7;  Conservative  7;  Mismatches  7;  Indels  0;  Gaps  0;

Qy   12 SCSVLDSFGELIPQPSN 28
         :::||||:|:|||:::|  :|
Db  153 SSKIMESFGKMVPQKAN 169
```

RESULT 3

```
ID   Q81224
AC   Q81224;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EXTRA-LARGE G-PROTEIN.
GN   XLG.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC   core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC   Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   LEE Y.-R.J., ASSMANN S.M.;
RT   "Arabidopsis thaliana extra-large G-protein: a new class of GTP
RT   binding protein.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF060941; AAC19352.1; -.
DR   HSSP; P10824; 1BH2.
DR   PFAM; PF00503; G-alpha; 1.
SQ   SEQUENCE 888 AA; 98796 MW; 46255612 CRC32;

Query Match    33.7%;  Score 49.5;  DB 10;  Length 888;
Best Local Similarity   43.6%;
Matches   17;  Conservative  3;  Mismatches  8;  Indels  11;  Gaps  3;

Qy   2 DCQLSI---LLLLSCSVLDSF------GELIPQ-PSNE 29
         |:|:||  ||| |||| |||       ||||| ||
Db  116 DCELSSSGELLRSCSVKESLDLNESSSNPLVPDWESNE 154
```

RESULT 4

```
ID   Q81225
AC   Q81225;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EXTRA-LARGE G-PROTEIN.
GN   XLG.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC   core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC   Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG;
RA   LEE Y.-R.J., ASSMANN S.M.;
RT   "Arabidopsis thaliana extra-large G-protein: a new class of GTP
RT   binding protein.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF060942; AAC19353.1; -.
DR   HSSP; P10824; 1BH2.
DR   PFAM; PF00503; G-alpha; 1.
SQ   SEQUENCE 888 AA; 98860 MW; 3F6B542C CRC32;

Query Match    33.7%;  Score 49.5;  DB 10;  Length 888;
Best Local Similarity   43.6%;
Matches   17;  Conservative  3;  Mismatches  8;  Indels  11;  Gaps  3;

Qy   2 DCQLSI---LLLLSCSVLDSF------GELIPQ-PSNE 29
         |:|:||  ||| |||| |||       ||||| ||
Db  116 DCELSSSGELLRSCSVKESLDLNESSSNPLVPDWESNE 154
```

RESULT 5

```
ID   Q80462
AC   Q80462;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PUTATIVE GUANINE NUCLEOTIDE-BINDING PROTEIN.
GN   F26B6.11.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC   core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC   Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   ROUNSLEY S.D., KAUL S., LIN X., KETCHUM K.A., CROSBY M.L.,
RA   BRANDON R.C., SYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA   SOMERVILLE C.R., VENTER J.C.;
RT   "Arabidopsis thaliana chromosome II BAC F26B6 genomic sequence.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC003040; AAC23761.1; -.
DR   HSSP; P10824; 1AS2.
DR   PFAM; PF00503; G-alpha; 1.
SQ   SEQUENCE 901 AA; 100285 MW; 93E6ED24 CRC32;

Query Match    33.7%;  Score 49.5;  DB 10;  Length 901;
Best Local Similarity   43.6%;
Matches   17;  Conservative  3;  Mismatches  8;  Indels  11;  Gaps  3;

Qy   2 DCQLSI---LLLLSCSVLDSF------GELIPQ-PSNE 29
         |:|:||  ||| |||| |||       ||||| ||
Db  116 DCELSSSGELLRSCSVKESLDLNESSSNPLVPDWESNE 154
```

```
RESULT 6

ID   Q23683      PRELIMINARY;   PRT;   147 AA.
AC   Q23683;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   ZK970.7 PROTEIN.
GN   ZK970.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BERKS M.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z49073; CAA88891.1; -.
SQ   SEQUENCE   147 AA;  15913 MW;  37A83B86 CRC32;

Query Match             33.3%;      Score 49; DB 5; Length 147;
Best Local Similarity   37.5%;      Pred. No. 4.9;
Matches    9;  Conservative   7;  Mismatches   8;  Indels   0;  Gaps   0;

Qy   3  CQLSILILLSCSVLDSFGEELIPQP  26
         |:|:||||| :|:|| |
Db   2  CRLAIALFSVAMVSVYGQDPPP  25
```

```
RESULT 7

ID   Q9WTX3      PRELIMINARY;   PRT;   349 AA.
AC   Q9WTX3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOSINE DNA METHYLTRANSFERASE (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE OF 1-69 FROM N.A.
RX   MEDLINE; 98389705.
RA   DENG J., SZYF M.;
RT   "Multiple isoforms of DNA methyltransferase are encoded by the
RT   vertebrate cytosine DNA methyltransferase gene.";
RL   J. Biol. Chem. 273:22869-22872(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   DENG J., SZYF M.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF083043; AAD28102.1; -.
DR   EMBL; AF083038; AAD28102.1; JOINED.
DR   EMBL; AF083039; AAD28102.1; JOINED.
DR   EMBL; AF083040; AAD28102.1; JOINED.
DR   EMBL; AF083041; AAD28102.1; JOINED.
DR   EMBL; AF083042; AAD28102.1; JOINED.
KW   Transferase; Methyltransferase.
FT   NON_TER        1      1
FT   NON_TER      349    349
SQ   SEQUENCE   349 AA;  39021 MW;  CCB24515 CRC32;

Query Match             32.0%;      Score 47; DB 11; Length 349;
Best Local Similarity   33.3%;      Pred. No. 22;
Matches    8;  Conservative   8;  Mismatches   8;  Indels   0;  Gaps   0;

Qy   2  DCQLSILILLSCSVLDSFGEELIPQ  25
         | : :|: |:|| :||| | :||
Db  27  DCNVLIKLVMAGEVTNSLGQRLPQ  50
```

```
RESULT 8

ID   O02232      PRELIMINARY;   PRT;   399 AA.
AC   O02232;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   C54C8.4 PROTEIN.
GN   C54C8.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   DOBSON R.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RL   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z83102; CAB05466.1; -.
SQ   SEQUENCE   399 AA;  46244 MW;  E9F94D5C CRC32;

Query Match             32.0%;      Score 47; DB 5; Length 399;
Best Local Similarity   62.5%;      Pred. No. 25;
Matches   10;  Conservative   3;  Mismatches   3;  Indels   0;  Gaps   0;

Qy   3  CQLSILILLSCSVLDS  18
         |:|  |:|||||| :|
Db   6  CKLTIKILLSCFVIYS  21
```

```
RESULT 9

ID   O27520      PRELIMINARY;   PRT;   429 AA.
AC   O27520;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TRYPTOPHAN SYNTHASE, BETA SUBUNIT HOMOLOG.
GN   MTH1476.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacterium.
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DELTA H;
RX   MEDLINE; 98037514.
RA   SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA   ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA   HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA   SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA   JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA   MCDOUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA   DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
DR   EMBL; AE000908; AAB85951.1; -.
     HSSP; P00933; 2WSY.
DR   PFAM; PF00247; trp_syntB; 1.
SQ   SEQUENCE   429 AA;  47487 MW;  DC78318F CRC32;

Qy          4  QLSILLLLSCSVLD----------SF----------GELIPQSN  28
               |  | |:||:|          ||      |:||| |||
Db        137  QWGTALSLACSLMDLQCKVYMVRVSFNQKPPRKTIMQLYGGEVVPSPSN  185
```

Query Match
Best Local Similarity   32.0%;
Matches 14;   Conservative   5;   Mismatches   6;   Indels 24;   Gaps   2;
Score 47;   DB 1;   Length 429;

```
RESULT   10

ID   Q92VS8       PRELIMINARY;   PRT;   514 AA.
AC   Q92VS8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   F15K9.13 PROTEIN.
GN   F15K9.13.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC   core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC   Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   VYSOTSKAIA V.S., SCHWARTZ J.R., TORIUMI M., YU G., LI J., LIU S.,
RA   KREMENETSKAIA I., LUROS J., ARAUJO R., BUEHLER E., CONWAY A.B.,
RA   DEWER K., FENG J., KIM C., LI Y., SHINN P., SUN H., DAVIS R.W.,
RA   ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.;
RT   "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RL   THEOLOGIS A.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RL   THEOLOGIS A.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RL   THEOLOGIS;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005278; AAC72127.1; -.
DR   MENDEL; 38431; Arath;3293;38431.
SQ   SEQUENCE   514 AA;  56415 MW;  B6C24CB4 CRC32;
```

Query Match   32.0%;   Score 47;   DB 10;   Length 514;

Best Local Similarity   47.4%;
Matches   9;   Conservative   6;   Mismatches   4;   Indels   0;   Gaps   0;
Score 47;   Pred. No. 32;

```
Qy          7  ILILLLSCSVLDSFGELIPQ  25
               : ||| ||| : ||||:|||
Db        120  VAVLLSVTFVLAFGEIIPQ  138
```

```
RESULT   11

ID   Q9Y0Y9       PRELIMINARY;   PRT;   1146 AA.
AC   Q9Y0Y9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   (1)
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   RUBIN G.M., WAN K.H., HARVEY D., LEWIS S.E., BROKSTEIN P., TSANG G.,
RA   AGBAYANI A., ARCAINA T.T., BAXTER E., BLAZEJ R.G., BUTENHOFF C.,
RA   CHAMPE M., CHAVEZ C., CHEW M., DOYLE C.M., FARFAN D.E., FRISE E.,
RA   GALLE R., GEORGE R.A., HARRIS N.L., HOSKINS R.A., EVANS-HOLM M.,
RA   HOUSTON K.A., HUMMASTI S.R., KIM E., LI P., MOSHREFI M., PACLEB J.M.,
RA   PARK S., SEQUEIRA A., SETHI H., SNIR H., SVIRSKAS R.R., WEINBURG T.,
RA   CELNIKER S.E.;
RT   "Full length Drosophila melanogaster cDNA sequence.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF145687; AAD38662.1; -.
SQ   SEQUENCE   1146 AA;  124541 MW;  8DA1D02B CRC32;
```

Query Match
Best Local Similarity   32.0%;
Matches   9;   Conservative   1;   Mismatches   4;   Indels   0;   Gaps   0;
Score 47;   DB 5;   Length 1146;

```
Qy         16  LDSFGELIPQPSNE  29
               ||| || ||| |:
Db        330  LDSVGEKTPQPDNK  343
```

```
RESULT   12

ID   Q92330       PRELIMINARY;   PRT;   1622 AA.
AC   Q92330;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOSINE-SPECIFIC
DE   METHYLTRANSFERASE).
OS   Rattus rattus (black rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND PLACENTA;
RX   MEDLINE; 99097263.
RA   KIMURA H., TAKEDA T., TANAKA S., OGAWA T., SHIOTA K.;
RT   "Expression of rat DNA (cytosine-5) methyltransferase (DNA MTase) in
RT   rodent trophoblast giant cells: molecular cloning and characterization
RT   of rat DNA MTase.";
RL   Biochem. Biophys. Res. Commun. 253:495-501(1998).
CC   -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA = S-
CC       ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.
CC   -!- SIMILARITY: STRONG, TO OTHER C5-DNA METHYLASES.
DR   EMBL; AB012214; BAA37118.1; -.
DR   PROSITE; PS00094; C5_MTASE_1; 1.
DR   PROSITE; PS00095; C5_MTASE_2; 1.
KW   Transferase; Methyltransferase; Restriction system.
```

FT   ACT_SITE   1231
SQ   SEQUENCE   1622 AA; 183000 MW; 69397C4D CRC32;
                   BY SIMILARITY.

Qy   2 DCQLSILLLSCSVLDSFGELIPQ 25
        |:|:|:||| ||:|:|||:||
Db 1195 DCNVLLKLVMAGEVTNSLGQRLPQ 1218

RESULT 13
ID   094559   PRELIMINARY;   PRT;   431 AA.
AC   094559;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   [2]
DE   HYPOTHETICAL 49.5 KD PROTEIN.
GN   SPAC23C4.05C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Archiascomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   HARRIS D., SQUARES R.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   BARRELL B.G., RAJANDREAM M.A., WOOD V.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99753; CAB16876.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   431 AA; 49532 MW; 67F4D4CF CRC32;

Query Match   31.3%; Score 46; DB 3; Length 431;
Best Local Similarity   72.7%; Pred. No. 38;
Matches   8; Conservative   2; Mismatches   1; Indels   0; Gaps   0;

Qy   16 LDSFGELIPQP 26
        |:||||||:| |
Db 375 LESFGELVPVP 385

RESULT 14
ID   018215   PRELIMINARY;   PRT;   793 AA.
AC   018215;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   Y53C12B.1 PROTEIN.
GN   Y53C12B.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditida; Rhabditina;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   KERSHAW J., LENNARD N.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z99278; CAB16490.1; -.
DR   PFAM; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINRPT.
SQ   SEQUENCE   793 AA; 88060 MW; E276DC23 CRC32;

Query Match   31.3%; Score 46; DB 5; Length 793;
Best Local Similarity   34.0%; Pred. No. 69;
Matches   16; Conservative   5; Mismatches   8; Indels 18; Gaps   3;

Qy   1 MDCQL------SIL-------LILSCSVLDS--FGELIPQPSNE 29
        |:|||| |:| |||| ||| |:| |||:|:|
Db 351 LDCQLIRGHTESVLSVVSPTWDTSLLASCSKDNSIIFWRLVTSPEND 397

RESULT 15
ID   Q88893   PRELIMINARY;   PRT; 2304 AA.
AC   Q88893;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   RNA1 POLYPROTEIN.
OS   Tobacco ringspot virus (Tobrsv) (Trsv).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC   Nepovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BUD BLIGHT;
RX   MEDLINE; 96204566.
RA   ZALLOUA P.A., BUZAYAN J.M., BRUENING G.;
RT   "Chemical cleavage of 5'-linked protein from tobacco ringspot virus
RT   genomic RNAs and characterization of the protein-RNA linkage.";
RL   Virology 219;1-8(1996).
DR   EMBL; U50869; AAB03785.1; -.
DR   PFAM; PF00680; RNA_dep_RNA_pol; 1.
DR   PFAM; PF00910; RNA_helicase; 1.
KW   Polyprotein
SQ   SEQUENCE   2304 AA; 256615 MW; 529C30B3 CRC32;

Query Match   31.3%; Score 46; DB 12; Length 2304;
Best Local Similarity   50.0%; Pred. No. 1.9e+02;
Matches   7; Conservative   5; Mismatches   2; Indels   0; Gaps   0;

Qy   10 LLSCSVLDSFGELI 23
        |:|| ||| |:|:||
Db 1775 LMSCQVLDAIGKMV 1788

Search completed: May   9, 2000, 22:31:13
Job time: 2378 sec

```
OM protein - protein search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:       May  9, 2000, 22:32:08 ; Search time 49.21 Seconds
                                        (without alignments)
                                        17.947 Million cell updates/sec

Title:
Perfect score:      147
Sequence:           US-09-104-340-3
Sequence:           1 MDCQLSILLLSCSVLDSFGELIPQPSNE 29

Post-processing: Minimum Match 0%
Maximum DB seq length: 1000000
Minimum DB seq length: 0

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:           83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters:     83857

Database :          SwissProt_38:*
                    Listing first 45 summaries

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	100.0	983	1	EPA3_HUMAN	P29320 homo sapien
2	105	71.4	983	1	EPA3_MOUSE	P29319 mus musculu
3	94	63.9	984	1	EPA3_RAT	O08680 rattus norv
4	52.5	35.7	500	1	CAMA_MOUSE	P51942 mus musculu
5	51	34.7	224	1	YO91_NPVAC	P41179 autographa
6	51	34.7	983	1	EPA3_CHICK,	P29318 gallus gall
7	47	32.0	1620	1	MTDM_MOUSE	P13864 mus musculu
8	46	31.3	189	1	INA1_PIG	P49079 sus scrofa
9	46	31.3	529	1	TDT_MOUSE	P09838 mus musculu
10	45	30.6	306	1	FAD1_YEAST	P38913 saccharomyc
11	45	30.6	370	1	PSPB_RABIT	P15285 oryctolagus
12	45	30.6	781	1	DPOL_ARCFU	O29753 archaeoglob
13	45	30.6	810	1	SYFB_SYNY3	P74296 synechocyst
14	45	30.6	908	1	MGR8_HUMAN	O00222 homo sapien
15	45	30.6	438	1	FIBG_XENLA	P17634 xenopus lae
16	44.5	30.3	1039	1	ATHL_HUMAN	P54707 homo sapien
17	44.5	30.3	80	1	CALT_BOVIN	P06833 bos taurus
18	43	29.3	585	1	PTRR_DIDMA	P25107 didelphis m
19	43	29.3	3801	1	LYST_HUMAN	Q99698 homo sapien
20	42.5	28.9	509	1	CPVI_BRARE	O42145 brachydanio
21	42.5	28.9	575	1	YFBK_ECOLI	P76481 escherichia
22	42.5	28.9	716	1	RRP2_IAGU2	P11167 influenza a
23	42	28.6	125	1	NU1M_ARBLI	Q33756 arbacia lix
24	42	28.6	189	1	INA1_BOVIN	P07348 bos taurus
25	42	28.6	189	1	INAB_BOVIN	P05008 bos taurus
26	42	28.6	189	1	INAD_BOVIN	P05010 bos taurus
27	42	28.6	189	1	INAF_BOVIN	P49876 bos taurus
28	42	28.6	189	1	INAG_BOVIN	P49877 bos taurus
29	42	28.6	189	1	INAH_BOVIN	P49878 bos taurus
30	42	28.6	234	1	GU45_RAT	P35898 rattus norv
31	42	28.6	257	1	FENR_AZOVI	Q44532 azotobacter
32	42	28.6	322	1	NU1M_STRPU	P15048 strongyloce
33	42	28.6	331	1	YOFA_BACSU	P54466 bacillus su
34	42	28.6	340	1	GLN2_STRHY	P22878 streptomyce
35	42	28.6	345	1	LEU3_THETH	P00351 thermus aqu
36	42	28.6	376	1	PSPB_RAT	P22355 rattus norv
37	42	28.6	558	1	AGP3_YEAST	P43548 saccharomyc
38	42	28.6	855	1	ENV_FIVT2	O02282 feline lmmu
39	42	28.6	908	1	MGR8_MOUSE	P47743 mus musculu
40	42	28.6	908	1	MGR8_RAT	P70579 rattus norv
41	42	28.6	1071	1	PR16_YEAST	P15938 saccharomyc
42	42	28.6	1133	1	SRE1_CRIGR	Q60416 cricetulus
43	41.5	28.2	246	1	MOG_MOUSE	Q61885 mus musculu
44	41.5	28.2	351	1	WNT4_MOUSE	P22724 mus musculu
45	41.5	28.2	648	1	TALA_POVMK	P24597 mouse polyo

RESULT 1
EPA3_HUMAN

ALIGNMENTS

```
ID   EPA3_HUMAN     STANDARD;      PRT;   983 AA.
AC   P29320;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (HEK).
GN   EPHA3 OR ETK1 OR ETK OR HEK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92179233.
RA   Wicks I.P., Wilkinson D., Salvaris E., Boyd A.W.;
RT   "Molecular cloning of HEK, the gene encoding a receptor tyrosine
RT   kinase expressed by human lymphoid tumor cell lines.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1611-1615(1992).
RN   [2]
RP   SEQUENCE OF 21-29 AND 840-860, AND CHARACTERIZATION.
RX   MEDLINE; 92147581.
RA   Boyd A.W., Ward L.D., Wicks I.P., Simpson R.J., Salvaris E., Wilks A.,
RA   Welch K., Loudovaris M., Rockman S., Busmanis I.;
RT   "Isolation and characterization of a novel receptor-type protein
RT   tyrosine kinase (hek) from a human pre-B cell line.";
RL   J. Biol. Chem. 267:3262-3267(1992).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5. COULD PLAY A ROLE IN LYMPHOID
CC       FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO LYMPHOID TUMOR CELL LINES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; M83941; AAA58633.1; -.
DR   EMBL; A28003; CAA01906.1; -.
DR   PIR; A38224; A38224.
DR   HSSP; P00523; 2PTK.
     MIM; 179611;
     PRINTS; PR00014; FNTYPEIII.
     PRINTS; PR00109; TYRKINASE.
     PRINTS; PR00107; TYRKINASE.
     PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
     PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
     PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
```

```
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL         1     20
FT   CHAIN         21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN        21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     542    565       POTENTIAL.
FT   DOMAIN       566    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        21    320       CYS-RICH.
FT   DOMAIN       321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN       432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN       621    882       PROTEIN KINASE.
FT   NP_BIND      627    635       ATP (BY SIMILARITY).
FT   BINDING      653    653       ATP (BY SIMILARITY).
FT   ACT_SITE     746    746       BY SIMILARITY.
FT   MOD_RES      779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     232    232       POTENTIAL.
FT   CARBOHYD     337    337       POTENTIAL.
FT   CARBOHYD     391    391       POTENTIAL.
FT   CARBOHYD     404    404       POTENTIAL.
FT   CARBOHYD     493    493       POTENTIAL.
FT   CONFLICT     507    507       F -> L (IN CAA01906).
FT   CONFLICT     724    724       V -> L (IN CAA01906).
SQ   SEQUENCE   983 AA;  110086 MW;  B8D900FA80FF5121 CRC64;

Query Match
Best Local Similarity  100.0%;      Score 147; DB 1; Length 983;
Matches   29; Conservative    0; Pred. No. 3.9e-14; Mismatches  0; Indels  0; Gaps  0;

Qy   1  MDCQLSILILLLSCSVLDSFGELIPQPSNE  29
        ||||||||||||||||||||||||||||||
Db   1  MDCQLSILILLLSCSVLDSFGELIPQPSNE  29

RESULT 2
ID   EPA3_MOUSE      STANDARD;      PRT;    983 AA.
AC   P29319;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (MEK4).
GN   EPHA3 OR ETK1 OR MEK4 OR TYRO4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR X SWISS WEBSTER; TISSUE=EMBRYO;
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SECRETED VARIANT WITHOUT THE TRANSMEMBRANE
CC       REGION AND THE KINASE DOMAIN IS PRODUCED BY ALTERNATIVE SPLICING
CC       OF THE SAME GENE.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURRING IN THE
CC       BRAIN.
```

```
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -----------------------------------------------------------------------
```

```
DR   EMBL; M68513; AAA39521.1; -.
DR   EMBL; M68515; AAA39522.1; ALT_SEQ.
DR   PIR; A45583; A45583.
DR   HSSP; P16109; 1FSB.
DR   MGD; MGI:99612; EPHA3.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL         1     20       BY SIMILARITY.
FT   CHAIN         21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN        21    540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     541    564       POTENTIAL.
FT   DOMAIN       565    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        21    320       CYS-RICH.
FT   DOMAIN       321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN       432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN       621    882       PROTEIN KINASE.
FT   NP_BIND      627    635       ATP (BY SIMILARITY).
FT   BINDING      653    653       ATP (BY SIMILARITY).
FT   ACT_SITE     746    746       BY SIMILARITY.
FT   MOD_RES      779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     231    231       POTENTIAL.
FT   CARBOHYD     336    336       POTENTIAL.
FT   CARBOHYD     390    390       POTENTIAL.
FT   CARBOHYD     403    403       POTENTIAL.
FT   CARBOHYD     492    492       POTENTIAL.
FT   VARSPLIC     530    983       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   983 AA;  109955 MW;  BE444A6655D8107A2 CRC64;

Query Match
Best Local Similarity  71.4%;      Score 105; DB 1; Length 983;
Matches   21; Conservative    1;  72.4%; Pred. No. 7.1e-08; Mismatches  7; Indels  0; Gaps  0;

Qy   1  MDCQLSILILLLSCSVLDSFGELIPQPSNE  29
        ||| |||||:||:||  |  ||| ||||||
Db   1  MDCHLSILVLLGCCVLSCSGGELSPQPSNE  29

RESULT 3
ID   EPA3_RAT        STANDARD;      PRT;    984 AA.
AC   O08680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR REK4).
GN   EPHA3 OR REK4.
OS   Rattus norvegicus (Rat).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE; 98120505.
RA   Li Y.Y., McIernan C.F., Feldman A.M.;
RT   "Il-1 beta alters the expression of the receptor tyrosine kinase gene
RT   r-EphA3 in neonatal rat cardiomyocytes.";
RL   Am. J. Physiol. 274:H331-H341(1998).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U69278; AAC06273.1; -.
DR   HSSP; P16109; 1B8B.
DR   PRINTS; PR00014; FN3TYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21    984       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    542    565
FT   DOMAIN      566    984       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    321       CYS-RICH.
FT   DOMAIN      322    432       FIBRONECTIN TYPE-III.
FT   DOMAIN      433    529       FIBRONECTIN TYPE-III.
FT   DOMAIN      622    883       PROTEIN KINASE.
FT   NP_BIND     628    636       ATP (BY SIMILARITY).
FT   BINDING     654    654       ATP (BY SIMILARITY).
FT   ACT_SITE    747    747       BY SIMILARITY.
FT   MOD_RES     780    780       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    232    232       POTENTIAL.
FT   CARBOHYD    337    337       POTENTIAL.
FT   CARBOHYD    391    391       POTENTIAL.
FT   CARBOHYD    404    404       POTENTIAL.
FT   CARBOHYD    493    493       POTENTIAL.
SQ   SEQUENCE   984 AA;  110227 MW;  F170C49312F7A0AB CRC64;

Qy   Query Match      63.9%;  Score 94;  DB 1;  Length 984;
     Best Local Similarity   65.5%;  Pred. No. 3.1e-06;
     Matches   19;  Conservative   1;  Mismatches   9;  Indels   0;  Gaps   0;

Qy   1   MDCQLSILILLSCSVLDSFGELIPQPSNE 29
         || ||||||:|:|| .|| || ||||||||
Db   1   MDCHLSILILFGCCVLSCSRELSPQPSNE 29
```

```
RESULT 4
CAMA_MOUSE
ID   CAMA_MOUSE     STANDARD;       PRT;   500 AA.
AC   P51942;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).
GN   MATN1 OR CRTM OR CMP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=CARTILAGE;
RX   MEDLINE; 96270751.
RA   Aszodi A., Hauser N., Studer D., Paulsson M., Hiripi L., Bosze Z.;
RT   "Cloning, sequencing and expression analysis of mouse cartilage
RT   matrix protein cDNA.";
RL   Eur. J. Biochem. 236:970-977(1996).
CC   -!- FUNCTION: CARTILAGE MATRIX PROTEIN IS A MAJOR COMPONENT OF THE
CC       EXTRACELLULAR MATRIX OF NONARTICULAR CARTILAGE. IT BINDS TO
CC       COLLAGEN.
CC   -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U35035; AAB05521.1; -.
DR   HSSP; P05099; 1AQ5.
DR   MGD; MGI:1095591; MATN1.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PFAM; PF00008; EGF; 1.
DR   PFAM; PF00092; vwa; 2.
KW   EGF-like domain; Signal; Glycoprotein; Cartilage; Repeat; Coiled coil.
FT   SIGNAL        1     29
FT   CHAIN        30    500       CARTILAGE MATRIX PROTEIN.
FT   DOMAIN       30    226       CMP 1.
FT   REPEAT      227    267       EGF-LIKE.
FT   DOMAIN      268    457       CMP 2.
FT   REPEAT      471    499       COILED COIL (POTENTIAL).
FT   DOMAIN       39    225       POTENTIAL.
FT   DISULFID    231    242       BY SIMILARITY.
FT   DISULFID    238    251       BY SIMILARITY.
FT   DISULFID    253    266       BY SIMILARITY.
FT   DISULFID    269    456       POTENTIAL.
FT   CARBOHYD     80     80       POTENTIAL.
FT   CARBOHYD    348    348       POTENTIAL.
SQ   SEQUENCE   500 AA;  54446 MW;  574602BF7E1EFFA6 CRC64;

Qy   Query Match      35.7%;  Score 52.5;  DB 1;  Length 500;
     Best Local Similarity   54.2%;  Pred. No. 2.3;
     Matches   13;  Conservative   2;  Mismatches   8;  Indels   1;  Gaps   1;

Qy   3   CQLSILILLSCSVLDSFGELIPQP 26
         || |||||:|||| .||| ||||||
Db   12  CSLLLLLLLQVPDSL-SLVPQP 34
```

```
RESULT 5
Y09I_NPVAC
ID   Y09I_NPVAC     STANDARD;       PRT;   224 AA.
AC   P41479;
DT   01-NOV-1995 (Rel. 32, Created)
```

```
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 24.1 KD PROTEIN IN LEF4-P33 INTERGENIC REGION.
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE; 94303173.
RA   Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT   "The complete DNA sequence of Autographa californica nuclear
RT   polyhedrosis virus.";
RL   Virology 202:586-605(1994).
CC   -!- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L22858; AAA66721.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   224 AA;  24138 MW;  13C016A64FD6A645 CRC64;

Query Match             34.7%;  Score 51;  DB 1;  Length 224;
Best Local Similarity   45.5%;
Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      5 LSILLLLSCSVLDSFGELIPQP 26
            | |:|||:|| :|| |||||
Db     16 LFIIILLTLKILDDIQELYPNP 37


RESULT 6
ID   EPA3_CHICK     STANDARD;      PRT;   983 AA.
AC   P29318;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (CEK4).
GN   EPHA3 OR ETK1 OR CEK4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       A PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURRING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M68514; AAA18666.1; -.
DR   PIR; B45583; B45583.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; 2PTK.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     19       BY SIMILARITY.
FT   CHAIN        20    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       20    540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    541    564       POTENTIAL.
FT   DOMAIN      565    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      321    320       CYS-RICH.
FT   DOMAIN      321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN      432    528       FIBRONECTIN TYPE-III.
FT   NP_BIND     621    882       PROTEIN KINASE.
FT   BINDING     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    231    231       POTENTIAL.
FT   CARBOHYD    336    336       POTENTIAL.
FT   CARBOHYD    390    390       POTENTIAL.
FT   CARBOHYD    403    403       POTENTIAL.
FT   CARBOHYD    492    492       POTENTIAL.
SQ   SEQUENCE   983 AA;  109910 MW;  E8895F0BDF77651E CRC64;

Query Match             34.7%;  Score 51;  DB 1;  Length 983;
Best Local Similarity   52.4%;
Matches   11;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      9 LILSCSVLDSFGELIPQPSNE 29
            |||| ||| |||||
Db      8 LILLCAALSAGRLSARPGNE 28


RESULT 7
ID   MTDM_MOUSE     STANDARD;      PRT;  1620 AA.
AC   P13864;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   DNA (CYTOSINE-5)-METHYLTRANSFERASE (EC 2.1.1.37) (DNA
DE   METHYLTRANSFERASE) (DNA METASE) (MCMT) (M.MMUI).
GN   DNMT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89094873.
RA   Bestor T.H., Laudano A., Mattaliano R., Ingram V.;
RT   "Cloning and sequencing of a cDNA encoding DNA methyltransferase of
RT   mouse cells. The carboxyl-terminal domain of the mammalian enzymes is
RT   related to bacterial restriction methyltransferases.";
RL   J. Mol. Biol. 203:971-983(1988).
```

RN   [2]
RP   REVISIONS TO N-TERMINUS.
RX   MEDLINE; 97094871.
RA   Yoder J.A., Vertino P.M., Bestor T.H., Baylin S.B.;
RT   "New 5' regions of the murine and human genes for DNA (cytosine-5)-
RT   methyltransferase.";
RL   J. Biol. Chem. 271:31092-31097(1996).
CC   -!- FUNCTION: METHYLATES CG RESIDUES.
CC   -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA = S-ADENOSYL-L-
CC   HOMOCYSTEINE + DNA CONTAINING 5-METHYLCYTOSINE.
CC   -!- SIMILARITY: HIGH, TO OTHER EUKARYOTIC DNA METASE.
CC   -!- SIMILARITY: SOME TO BACTERIAL RESTRICTION SYSTEMS
CC   METHYLTRANSFERASES.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; X14805; CAA32910.1; -.
DR   PIR; S01845; S01845.
DR   REBASE; RB02813; M.Mmul.
DR   MGD; MGI:94912; Dnmt.
DR   PRINTS; PR00105; C5METHTRFRASE.
DR   PROSITE; PS00094; C5_MTASE_1; 1.
DR   PROSITE; PS00095; C5_MTASE_2; 1.
DR   PFAM; PF00145; DNA_methylase; 3.
DR   PFAM; PF01426; BAH; 2.
KW   Transferase; Methyltransferase; DNA-binding.
FT   DOMAIN   651   693   CYS/ARG/LYS-RICH.
FT   ACT_SITE   1229   1229   BY SIMILARITY.
SQ   SEQUENCE   1620 AA; 183286 MW; F73100AD043E709 CRC64;

Qy   2 DCQLSILILLSCSVLDSFGELIPQ 25
             |::| |:|:| |:| | ||:||
Db 1193 DCNVLKLVMAGEVTNSLGQRLPQ 1216

Query Match   32.0%; Score 47; DB 1; Length 1620;
Best Local Similarity   33.3%;
Matches   8; Conservative   8; Mismatches 51; Pred. No. 8; Indels 0; Gaps 0;

RESULT   8
ID   INA1_PIG   STANDARD;   PRT;   189 AA.
AC   P49879;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   INTERFERON ALPHA-1 PRECURSOR (IFN-ALPHA-1).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DOMESTICA; TISSUE=LIVER;
RX   MEDLINE; 92193689.
RA   Lefevre F., la Bonnardiere C., Mege D.;
RT   "The porcine family of interferon-omega: cloning, structural
RT   analysis, and functional studies of five related genes.";
RL   J. Interferon Res. 11:341-350(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87035519.
RA   Lefevre F., la Bonnardiere C.;
RT   "Molecular cloning and sequencing of a gene encoding biologically
RT   active porcine alpha-interferon.";
RL   J. Interferon Res. 6:349-360(1986).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL

CC   ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC   A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC   FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; X57191; CAA40477.1; -.
DR   EMBL; M28623; AAA31053.1; -.
DR   HSSP; P01563; 1ITF.
DR   PRINTS; PR00266; INTERFERONAB.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
DR   PFAM; PF00143; interferon; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL   1   23   POTENTIAL.
FT   CHAIN   24   189   INTERFERON ALPHA-1.
FT   DISULFID   24   122   BY SIMILARITY.
FT   DISULFID   52   162   BY SIMILARITY.
SQ   SEQUENCE   189 AA; 21433 MW; 8C8BEA8DEAAE3334 CRC64;

Qy   5 LSILILLSCSVLDSFGELIPQ 25
            |:|:||||:| | || |:||
Db   8 LTALVLLSCNAICSLGCDLPQ 28

Query Match   31.3%; Score 46; DB 1; Length 189;
Best Local Similarity   47.6%;
Matches   10; Conservative   5; Mismatches 7; Pred. No. 7.9; Indels 0; Gaps 0;

RESULT   9
ID   TDT_MOUSE   STANDARD;   PRT;   529 AA.
AC   P09838;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   DNA NUCLEOTIDYLEXOTRANSFERASE (EC 2.7.7.31) (TERMINAL ADDITION ENZYME)
DE   (TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE) (TDT) (TERMINAL TRANSFERASE).
GN   DNTT OR TDT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=THYMUS;
RX   MEDLINE; 93219079.
RA   Doyen N., Fanton D'Andon M., Bentolila L.A., Nguyen T.Q., Rougeon F.;
RT   "Differential splicing in mouse thymus generates two forms of
RT   terminal deoxynucleotidyl transferase.";
RL   Nucleic Acids Res. 21:1187-1191(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Koiwai O., Yokota T., Kageyama T., Hirose T., Yoshida S., Arai K.;
RT   "Isolation and characterization of bovine and mouse terminal
RT   deoxynucleotidyltransferase cDNAs expressible in mammalian cells.";
RL   Nucleic Acids Res. 14:5777-5792(1986).
RN   [3]
RP   CHARACTERIZATION OF ALTERNATIVE FORMS.
RX   MEDLINE; 96016194.
RA   Bentolila L.A., Fanton D'Andon M., Nguyen T.Q., Martinez O.,
RA   Rougeon F., Doyen N.;
RT   "The two isoforms of mouse terminal deoxynucleotidyl transferase
RT   differ in both the ability to add N regions and subcellular
RT   localization.";
RL   EMBO J. 14:4221-4229(1995).

```
CC  -!- FUNCTION: TEMPLATE-INDEPENDENT DNA POLYMERASE WHICH CATALYZES THE
CC      RANDOM ADDITION OF DEOXYNUCLEOSIDE 5'-TRIPHOSPHATE TO THE 3' END OF
CC      A DNA INITIATOR. ONE OF THE IN-VIVO FUNCTION OF THIS ENZYME IS THE
CC      ADDITION OF NUCLEOTIDES AT THE JUNCTION (N REGION) OF REARRANGED
CC      IG HEAVY CHAIN AND T CELL RECEPTOR GENE SEGMENTS DURING THE
CC      MATURATION OF B AND T CELLS.
CC  -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE + (DEOXY
CC      NUCLEOTIDE)(M) = N PYROPHOSPHATE + (DEOXYNUCLEOTIDE)(M+N).
CC  -!- COFACTOR: REQUIRES MAGNESIUM.
CC  -!- SUBCELLULAR LOCATION: NUCLEAR (TDT-S) OR CYTOPLASMIC (TDT-L).
CC  -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS; TDT-L(ARGE) (SHOWN HERE) AND
CC      TDT-S(MALL); ARE PRODUCED BY ALTERNATIVE SPLICING. THE TDT-S FORM
CC      IS THE MAJOR FORM. THE TWO FORMS DIFFER IN SUBCELLULAR LOCATION
CC      AND IN ACTIVITY AS THE LONG CYTOPLASMIC FORM CAN NOT ACT ON N
CC      REGIONS.
CC  -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-X FAMILY.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR  EMBL; X04123; CAA27735.1; -.
DR  EMBL; X68670; CAA48634.1; -.
DR  PIR; B23595; B23595.
MGD  MGI:98659; TDT.
PRINTS; PR00869; DNAPOLX.
PRINTS; PR00871; DNAPOLXTDT.
PROSITE; PS00522; DNA_POLYMERASE_X; 1.
PFAM; PF00533; BRCT; 1.
PFAM; PF00966; DNA_polymeraseX; 1.
Transferase; Nucleotidyltransferase; Terminal addition; Magnesium;
Alternative splicing; Nuclear protein.
VARSPLIC  482  501  MISSING (IN ISOFORM TDT-S).
CONFLICT   26   26  M -> T (IN REF. 1).
CONFLICT   99   99  F -> L (IN REF. 1).
CONFLICT  193  193  R -> G (IN REF. 1).
CONFLICT  287  287  Q -> K (IN REF. 1).
CONFLICT  309  309  E -> Q (IN REF. 1).
CONFLICT  367  367  D -> H (IN REF. 1).
CONFLICT  441  444  DRAS -> ECAC (IN REF. 1).
SEQUENCE  529 AA;  60179 MW;  DE949FD7CE3A9562 CRC64;

Qy    9  LILLSCSVLDSFGELIPQPSNE  29
         |||  |:||:|    ||| :
Db  374  LILYCDILESTFEKFKQPSRK  394

Query Match             31.3%;
Best Local Similarity   47.6%;   Score 46;  DB 1;  Length 529;
Matches  10;  Conservative  3;  Mismatches  8;  Indels  0;  Gaps  0;

RESULT 10
FAD1_YEAST

ID  FAD1_YEAST     STANDARD;      PRT;   306 AA.
AC  P38913;
DT  01-FEB-1995 (Rel. 31, Created)
DT  01-FEB-1995 (Rel. 31, Last sequence update)
DT  15-FEB-2000 (Rel. 39, Last annotation update)
DE  FAD SYNTHETASE (EC 2.7.7.2) (FMN ADENYLYLTRANSFERASE) (FAD
DE  PYROPHOSPHORYLASE) (FLAVIN ADENINE DINUCLEOTIDE SYNTHETASE).
GN  FAD1 OR YDL045C OR D2702.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomyces; Saccharomycetales;
OC  Saccharomycetaceae; Saccharomyces.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=W303-1A / D273-10B;
RX  MEDLINE; 95098000.
RA  Wu M., Repetto B., Glerum D.M., Tzagoloff A.;
RT  "Cloning and characterization of FAD1, the structural gene for flavin
RT  adenine dinucleotide synthetase of Saccharomyces cerevisiae.";
RL  Mol. Cell. Biol. 15:264-271(1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S288C;
RA  Paulin L., Saren A.M., Laamanen P.;
RL  Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC  -!- FUNCTION: ADENYLATES FMN TO FAD.
CC  -!- CATALYTIC ACTIVITY: ATP + FMN = DIPHOSPHATE + FAD.
CC  -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC  -!- SIMILARITY: SOME, TO BACTERIAL SULFATE ADENYLATE TRANSFERASES.
CC  --------------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the European
CC  Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------------
DR  EMBL; U12331; AAA65730.1; -.
DR  EMBL; Z71781; CAA96644.1; -.
DR  EMBL; Z74093; CAA98604.1; -.
DR  PIR; S47906; S47906.
DR  SGD; L0000598; FAD1.
DR  PFAM; PF01507; PAPS_reduct; 1.
KW  Transferase; Nucleotidyltransferase; FAD; FMN.
SQ  SEQUENCE  306 AA;  35546 MW;  55BBB83016 3A457F CRC64;

Qy    2  DCQLSILILLSC  13
         ||| :|| |||
Db   65  DCQVLILIYLSC  76

Query Match             30.6%;
Best Local Similarity   66.7%;   Score 45;  DB 1;  Length 306;
Matches   8;  Conservative  2;  Mismatches  2;  Indels  0;  Gaps  0;

RESULT 11
PSPB_RABIT

ID  PSPB_RABIT     STANDARD;      PRT;   370 AA.
AC  P15285;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  PULMONARY SURFACTANT-ASSOCIATED PROTEIN B PRECURSOR (SP-B) (6 KD
DE  PROTEIN) (PULMONARY SURFACTANT-ASSOCIATED PROTEOLIPID SPL(PHE)).
GN  SFTPB OR SFTP3.
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=LUNG;
RX  MEDLINE; 89228033.
RA  Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,
RA  Hammond G., Possmayer F.;
RT  "Isolation and characterization of the cDNA for pulmonary surfactant-
RT  associated protein-B (SP-B) in the rabbit";
RL  Biochem. Biophys. Res. Commun. 160:325-332(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NEW ZEALAND WHITE;
RX  MEDLINE; 95208794.
RA  Margana R.K., Boggaram V.;
RT  "Transcription and mRNA stability regulate developmental and hormonal
RT  expression of rabbit surfactant protein B gene.";
RL  Am. J. Physiol. 268:L481-L490(1995).
RN  [3]
RP  SEQUENCE OF 1-34 FROM N.A.
```

RX   MEDLINE; 96965536.
RA   Luzi P., Anceschi M., Strayer D.S.;
RT   "The upstream region of the SP-B gene: Intrinsic promoter activity
RT   and glucocorticoid responsiveness related to a new DNA-binding
RT   protein.";
RL   Gene 165:285-290(1995).
CC   -!- FUNCTION: PULMONARY SURFACTANT ASSOCIATED PROTEINS PROMOTE
CC   ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC   LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC   THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC   PER METER.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC   PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CC   CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC   HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M24901; AAA31466.1; -.
DR   EMBL; U11106; AAA67934.1; -.
DR   EMBL; S80649; AAD14335.1; -.
DR   PIR; A32421; LNRBB.
KW   Surface film; Gaseous exchange.
FT   PROPEP   1   184
FT   CHAIN   185   263   PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT   B.
FT   DOMAIN   29   62   SAPOSINS-LIKE TYPE A.
FT   DISULFID   192   261   BY SIMILARITY.
FT   DISULFID   195   255   BY SIMILARITY.
FT   DISULFID   219   230   BY SIMILARITY.
FT   DISULFID   232   232   BY SIMILARITY.
FT   CARBOHYD   300   300   INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD   15   15   POTENTIAL.
FT   CONFLICT   15   15   MISSING (IN REF. 2).
FT   CONFLICT   184   184   R -> P (IN REF. 1).
FT   CONFLICT   232   232   C -> R (IN REF. 2).
FT   CONFLICT   329   355   ELHTPQLLSLLSRGWDARAICQALGAC ->
FT   AAHAPAAEPAVQGLGCPRNLPGPEGRV (IN REF. 1).
SQ   SEQUENCE   370 AA; 40610 MW; 423047A69B12DCB5 CRC64;

Query Match   30.6%; Score 45; DB 1; Length 370;
Best Local Similarity   47.6%;
Matches   10; Conservative   4; Mismatches   7; Indels   0; Gaps   0;

Qy   4 QLSILILLSCSVLDSFGELIP 24
              || ||:|| || :|||| |:|
Db   252 QLVCGLVLRCSSVDSIGQVPP 272

RESULT 12
ID   DPOL_ARCFU   STANDARD;   PRT;   781 AA.
AC   O29753;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DNA POLYMERASE (EC 2.7.7.7).
GN   POL OR POLB OR AF0497.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE; 98049343.

RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC   N PYROPHOSPHATE + DNA(N).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AE001070; AAB90741.1; -.
DR   TIGR; AF0497; -.
DR   PRINTS; PR00106; DNAPOLB.
DR   PROSITE; PS00116; DNA_POL_B; 1.
DR   PFAM; PF00136; DNA_pol_B; 1.
KW   Transferase; DNA-directed DNA polymerase; DNA replication;
KW   DNA-binding.
SQ   SEQUENCE   781 AA; 89850 MW; 761C7B18FCC61B8B CRC64;

Query Match   30.6%; Score 45; DB 1; Length 781;
Best Local Similarity   36.4%; Pred. No. 48;
Matches   8; Conservative   6; Mismatches   6; Indels   2; Gaps   1;

Qy   8 LILLSCSVLDSFGELIPQPSNE 29
              :|: || ||| :|:| :
Db   164 MLVFDCEMLSSFG--MPEPEKD 183

RESULT 13
ID   SYFB_SYNY3   STANDARD;   PRT;   810 AA.
AC   P74296;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (EC 6.1.1.20) (PHENYLALANINE--
DE   TRNA LIGASE BETA CHAIN) (PHERS).
GN   PHET OR SLL1553.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97061201.
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- CATALYTIC ACTIVITY: ATP + L-PHENYLALANINE + TRNA(PHE) = AMP +
CC   PYROPHOSPHATE + L-PHENYLALANYL-TRNA(PHE).
CC   -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS
CC   (BY SIMILARITY).

(Left column — continuation of previous RESULT)

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

DR EMBL; D90913; BAA18390.1; -.
DR HSSP; P27002; 1PYS.
DR PFAM; PF01588; tRNA_bind; 2.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 810 AA; 87887 MW; 4D14168774C12BE9 CRC64;

Query Match   54.5%; Score 45; DB 1; Length 810;
Best Local Similarity   30.6%; Pred. No. 50;
Matches   12; Conservative   2; Mismatches   6; Indels   2; Gaps   1;

Qy   10 LLSC--SVLDSFGELIPQPSNE 29
             ||| :||| ||||||||
Db 660 LLSCNGTVLGRFGOLHPOLRRE 681

RESULT 14

MGR8_HUMAN

ID MGR8_HUMAN   STANDARD;   PRT;   908 AA.
AC 000222; O15493;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE METABOTROPIC GLUTAMATE RECEPTOR 8 PRECURSOR.
GN GRM8 OR MGLUR8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98141892.
RA Scherer S.W., Soder S., Duvoisin R.M., Huizenga J.J., Tsui L.C.;
RT "The human metabotropic glutamate receptor 8 (GRM8) gene: a
RT disproportionately large gene located at 7q31.3-q32.1.";
RL Genomics 44:232-236(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Wu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S.,
RA Rosteck P.R.Jr., Johnson B.G., Schoepp D.D., Belagaje R.M.;
RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
RT molecular cloning, functional expression, and comparison of
RT pharmacological properties in RGT cells.";
RL Brain Res. Mol. Brain Res. 53:88-97(1998).
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC   IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC   ACTIVITY.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.

DR EMBL; U92459; AAB51764.1; -.
DR EMBL; U95025; AAB72040.1; -.
DR GCRDB; GCR_1889; -.
DR GCRDB; GCR_2604; -.
DR MIM; 601116; -.
DR PRINTS; PR00248; GPCRMGR.
DR PRINTS; PR00593; MTABOTROPICR.
DR PRINTS; PR01058; MTABOTROPC8R.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.
DR PFAM; PF00003; 7tm_3; 1.
DR PFAM; PF01094; ANF_receptor; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family; Olfaction.
FT SIGNAL   1   33   POTENTIAL.
FT DOMAIN   34   908   METABOTROPIC GLUTAMATE RECEPTOR 8.
FT CHAIN   34   583   EXTRACELLULAR (POTENTIAL).
FT DOMAIN   584   608   I (POTENTIAL).
FT TRANSMEM   609   620   CYTOPLASMIC (POTENTIAL).
FT TRANSMEM   621   641   II (POTENTIAL).
FT DOMAIN   642   647   EXTRACELLULAR (POTENTIAL).
FT TRANSMEM   648   668   III (POTENTIAL).
FT DOMAIN   669   695   CYTOPLASMIC (POTENTIAL).
FT TRANSMEM   696   716   IV (POTENTIAL).
FT TRANSMEM   717   746   EXTRACELLULAR (POTENTIAL).
FT DOMAIN   747   768   V (POTENTIAL).
FT DOMAIN   769   781   CYTOPLASMIC (POTENTIAL).
FT TRANSMEM   782   803   VI (POTENTIAL).
FT DOMAIN   804   818   EXTRACELLULAR (POTENTIAL).
FT TRANSMEM   819   843   VII (POTENTIAL).
FT DOMAIN   844   908   CYTOPLASMIC (POTENTIAL).
FT CARBOHYD   95   95   POTENTIAL.
FT CARBOHYD   298   298   POTENTIAL.
FT CARBOHYD   452   452   POTENTIAL.
FT CARBOHYD   480   480   POTENTIAL.
FT CARBOHYD   565   565   POTENTIAL.
FT CARBOHYD   194   194   POTENTIAL.
FT CARBOHYD   460   460   POTENTIAL.
FT CARBOHYD   642   642   POTENTIAL.
FT CARBOHYD   768   768   POTENTIAL.
FT CARBOHYD   904   904   POTENTIAL.
FT CONFLICT   34   34   S -> T (IN REF. 2).
FT CONFLICT   95   95   N -> I (IN REF. 2).
FT CONFLICT   298   298   A -> V (IN REF. 2).
FT CONFLICT   452   452   T -> A (IN REF. 2).
FT CONFLICT   460   460   R -> A (IN REF. 2).
FT CONFLICT   565   565   A -> V (IN REF. 2).
FT CONFLICT   642   642   A -> T (IN REF. 2).
FT CONFLICT   768   768   I -> V (IN REF. 2).
SQ SEQUENCE 908 AA; 101741 MW; 95C2D0886E743905 CRC64;

Query Match   37.9%; Score 45; DB 1; Length 908;
Best Local Similarity   30.6%; Pred. No. 56;
Matches   11; Conservative   7; Mismatches   9; Indels   2; Gaps   1;

Qy   3 CQL--SILLLLCGSVLDSFGELIPQPSNE 29
            |: ||||||||:||:| :|: ||
Db 751 CSLGYSILLMVTCTVYANKTRGVPETENE 779

RESULT 15

FIBG_XENLA

ID FIBG_XENLA   STANDARD;   PRT;   438 AA.
AC P17634;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBRINOGEN GAMMA CHAIN PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90241882.
RA Pastori R.L., Moskaitis J.E., Smith L.H. Jr., Schoenberg D.R.;
RT "Estrogen regulation of Xenopus laevis gamma-fibrinogen gene
RT expression.";
RL Biochemistry 29:2599-2605(1990).
RN [2]
RP SEQUENCE OF 1-58 FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 91146806.
RA Bhattacharya A., Shepard A.R., Moser D.R., Holland L.J.;

```
RT   "Isolation and characterization of cDNA clones for the gamma subunit
RT   of Xenopus fibrinogen, the product of a coordinately regulated gene
RT   family.";
RL   Mol. Cell. Endocrinol. 72:213-220(1990).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC       CONVERTED INTO THE HARD CLOT BY FACTOR XIIIA WHICH CATALYZES THE
CC       EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC       (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC       MONOMERS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J02894; AAA49709.1; -.
DR   EMBL; M35548; AAA03247.1; -.
DR   PIR; A32670; A32670.
DR   HSSP; P02679; 1FIC.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
DR   PFAM; PF00147; fibrinogen_C; 1.
KW   Blood coagulation; Glycoprotein; Calcium; Plasma; Platelet; Signal.
FT   SIGNAL        1     25
FT   CHAIN        26    438       FIBRINOGEN GAMMA CHAIN.
FT   DISULFID     32     32       INTERCHAIN (WITH C-33') (BY SIMILARITY).
FT   DISULFID     33     33       INTERCHAIN (WITH C-32') (BY SIMILARITY).
FT   DISULFID     43     43       INTERCHAIN (WITH BETA) (BY SIMILARITY).
FT   DISULFID     47     47       INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
FT   DISULFID    158    158       INTERCHAIN (WITH BETA) (BY SIMILARITY).
FT   DISULFID    162    162       INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
FT   DISULFID    176    205       BY SIMILARITY.
FT   DISULFID    349    362       BY SIMILARITY.
FT   CARBOHYD     76     76       POTENTIAL.
SQ   SEQUENCE   438 AA;  50064 MW;  69B1F01BB6716F60 CRC64;
```

```
Query Match             30.3%;   Score 44.5;  DB 1;  Length 438;
Best Local Similarity   47.8%;   Pred. No. 31;
Matches   11;  Conservative   4;  Mismatches   7;  Indels   1;  Gaps   1;

Qy   7 ILLLLSCSVLDS-FGELIPQPSN 28
       :||||  :::|||  ||  :|||
Db   9 LILLQSLALLSSAFGNIIPNTDN 31
```

Search completed: May 9, 2000, 22:32:11
Job time: 2402 sec

This Page Blank (uspto)

OM protein - protein search, using sw model

         GenCore version 4.5
         Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 22:27:16 ; Search time 64.29 Seconds
                                 (without alignments)
                                 26.447 Million cell updates/sec

Title: US-09-104-340-3
Perfect score: 147
Sequence: 1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29

Scoring table:
       Gapop 10.0 , Gapext 0.5
       BLOSUM62

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Maximum DB seq length: 1000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
           Listing first 45 summaries

Database :   PIR_63:*
          1:   pir1:*
          2:   pir2:*
          3:   pir3:*
          4:   pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	147	100.0	983	2	A38224	protein-tyrosine k
2	105	71.4	983	2	A45583	receptor tyrosine
3	58.5	39.8	495	2	T18835	hypothetical prote
4	52.5	35.7	500	2	S66522	cartilage matrix p
5	51	34.7	224	2	D72861	ACorf-91 protein -
6	51	34.7	983	2	B45583	receptor tyrosine
7	49.5	33.7	901	2	T01135	receptor tyrosine
8	49	33.3	147	2	T28131	hypothetical prote
9	47	32.0	399	2	T20204	hypothetical prote
10	47	32.0	429	2	F69063	tryptophan synthas
11	47	32.0	1622	2	JE0378	DNA (cytosine-5-)-
12	46	31.3	189	2	S23709	interferon alpha-1
13	46	31.3	431	2	T38260	hypothetical prote
14	46	31.3	529	2	B23595	DNA nucleotidylexo
15	46	31.3	793	2	T27133	hypothetical prote
16	45.5	31.0	526	2	T12313	permease 1 - commo
17	45.5	31.0	528	2	T12309	permease 1 - commo
18	45.5	31.0	975	2	T16073	hypothetical prote
19	45	30.6	79	2	S77428	hypothetical prote
20	45	30.6	223	2	T21437	hypothetical prote
21	45	30.6	369	2	S47906	FAD synthetase - y
22	45	30.6	369	2	I46531	surfactant protein
23	45	30.6	563	1	LNRBB	pulmonary surfacta
24	45	30.6	370	1	C64420	N-methylhydantoina
25	45	30.6	781	2	A69312	DNA polymerase B1
26	45	30.6	810	2	S75931	hypothetical prote
27	45	30.6	1490	2	JC5145	hypothetical prote
28	45	30.6	3898	2	S58295	DNA (cytosine-5-)-
29	44.5	30.3	128	2	S13028	polyprotein - hog
30	44.5	30.3	176	2	I41076	site-specific DNA-
31	44.5	30.3	276	2	S31504	H+/K+-exchanging A
32	44.5	30.3	438	2	A32670	fibrinogen gamma c
33	44.5	30.3	591	2	T39195	probable amino aci
34	44.5	30.3	1039	2	I38401	ATP-driven ion pum
35	44	29.9	111	2	D72112	hypothetical prote
36	44	29.9	157	2	S58068	probable olfactory
37	44	29.9	191	2	S67735	hypothetical prote
38	44	29.9	411	2	T04987	hypothetical prote
39	44	29.9	482	2	A70963	hypothetical prote
40	44	29.9	1337	2	JC4172	DNA (cytosine-5-)-
41	43.5	29.6	364	2	T10174	hypothetical prote
42	43	29.3	80	2	A35572	seminalplasmin pre
43	43	29.3	94	2	B45506	PRF2 protein - hum
44	43	29.3	117	2	T20888	hypothetical prote
45	43	29.3	130	1	JL0079	Ig kappa chain pre

ALIGNMENTS

RESULT 1
A38224
protein-tyrosine kinase (EC 2.7.1.112) hek precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C;Accession: A38224; B38224
R;Wicks, I.P.; Wilkinson, D.; Salvaris, E.; Boyd, A.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 1611-1615, 1992
A;Title: Molecular cloning of HEK, the gene encoding a receptor tyrosine kinase expre
A;Reference number: A38224; MUID:92179233
A;Accession: B38224
A;Molecule type: mRNA
A;Residues: 1-983 <WIC>
A;Cross-references: GB:M83941; NID:g183931; PIDN:AAA58633.1; PID:g183932
A;Experimental source: pre-B-cell leukemia cell line LK63
A;Note: sequence extracted from NCBI backbone (NCBIP:86627)
A;Molecule type: protein
A;Residues: 21-39;810-860 <WI2>
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferas
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-983/Product: protein-tyrosine kinase hek #status experimental <MAT>
F;542-565/Domain: transmembrane #status predicted <TMM>
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;232,337,391,404,493/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Qy   1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29
         |||||||||||||||||||||||||||||
Db   1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29
```

Query Match   100.0%;   Score 147;   DB 2;   Length 983;
Best Local Similarity   100.0%;   Pred. No. 1.2e-13;
Matches   29;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 2
A45583
receptor tyrosine kinase Mek4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse
A;Reference number: A45583; MUID:92031278
A;Accession: A45583
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68513; NID:g199119; PIDN:AAA39521.1; PID:g199120

A;Note: sequence extracted from NCBI backbone (NCBIN:62398, NCBIP:62401)
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Qy   1 MDCQLSILLLL-------------------------SCSVLDSFGELIPQPSNE 29
        ||| |||||||| |||||||||||||||||
Db   1 MDCHLSLLVLLGCCVLSCSGELSPQPSNE 29

RESULT 3
T18835
hypothetical protein C01G6.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999
C;Accession: T18835
R;Berks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: 219029
A;Accession: T18835
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-495 <WI>
A;Cross-references: EMBL:Z35595; PIDN:CAA84634.1; GSPDB:GN00020; CESP:C01G6.3
A;Experimental source: clone C01G6
C;Genetics:
A;Gene: CESP:C01G6.3
A;Map position: 2
A;Introns: 47/1; 303/3; 344/3

Query Match   71.4%; Score 105; DB 2; Length 983;
Best Local Similarity   72.4%; Pred. No. 1.8e-07;
Matches   21; Conservative   1; Mismatches   7; Indels   0; Gaps   0;

Qy   1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
        |||||| ||||||||||| |||||||||
Db   1 MDCHLSLLVLLGCCVLSCSGELSPQPSNE 29

RESULT 4
S66522
cartilage matrix protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C;Accession: S66522
R;Aszodi, A.; Hauser, N.; Studer, D.; Paulsson, M.; Hiripi, L.; Bosze, Z.
Eur. J. Biochem. 236, 970-977, 1996
A;Title: Cloning, sequencing and expression analysis of mouse cartilage matrix protein
A;Reference number: S66522; MUID:96270751
A;Accession: S66522
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-500 <ASZ>
A;Cross-references: EMBL:U35035; NID:g1163178; PIDN:AAB05521.1; PID:g1163179
C;Genetics:
A;Gene: CMP
C;Superfamily: von Willebrand factor type A repeat homology; EGF homology
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-500/Product: cartilage matrix protein #status predicted <MAT>
F;43-210/Domain: von Willebrand factor type A repeat homology <VWAl>
F;231-266/Domain: EGF homology <EGF>
F;277-441/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match   35.7%; Score 52.5; DB 2; Length 500;
Best Local Similarity   54.2%; Pred. No. 4.8;

Matches   13; Conservative   2; Mismatches   8; Indels   1; Gaps   1;

Qy   3 CQLSILLLLSCSVLDSFGELIPQP 26
        | ||||||||||||||||||||
Db   12 CSLLLLLLLLQVPDSL-SIVPQP 34

RESULT 5
D72861
AcOrf-91 protein - Autographa californica nuclear polyhedrosis virus
C;Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C;Accession: D72861
R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
A;Reference number: A72850; MUID:94303173
A;Accession: D72861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-224 <AYR>
A;Cross-references: GB:L22858; NID:g510708; PIDN:AAA66721.1; PID:g559160
C;Genetics:
A;Gene: AcOrf-91

Query Match   34.7%; Score 51; DB 2; Length 224;
Best Local Similarity   45.5%; Pred. No. 3.6;
Matches   10; Conservative   4; Mismatches   8; Indels   0; Gaps   0;

Qy   5 LSILLLLSCSVLDSFGELIPQP 26
        | ||||| ||||||||| |
Db   16 LFIIILILTLKILDDIQELYPNP 37

RESULT 6
B45583
receptor tyrosine kinase Cek4 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: B45583
A;Note: sequence extracted from NCBI backbone (NCBIN:62405, NCBIP:62411)
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse
A;Reference number: A45583; MUID:92031278
A;Accession: B45583
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   34.7%; Score 51; DB 2; Length 983;
Best Local Similarity   52.4%; Pred. No. 16;
Matches   11; Conservative   2; Mismatches   8; Indels   0; Gaps   0;

Qy   9 LLLLSCSVLDSFGELIPQPSNE 29
        ||| | || |||||| || |
Db   8 LILLCAALGSAGRLSARPGNE 28

RESULT 7
T01135
hypothetical protein F26B6.11 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 30-Apr-1999
C;Accession: T01135

R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F26B6 genomic sequence.
A;Reference number: 214198
A;Accession: T01135
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-901 <ROU>
A;Cross-references: EMBL:AC003040; NID:g3242700; PID:g3242709
C;Genetics:
A;Experimental source: cultivar Columbia
A;Map position: 2
A;Introns: 358/3; 439/1; 502/3; 565/1; 646/3; 701/2
A;Note: F26B6.11

Query Match   33.7%; Score 49.5; DB 2; Length 901;
Best Local Similarity 43.6%; Pred. No. 24;
Matches   17; Conservative   3; Mismatches   8; Indels 11; Gaps   3;

```
Qy   2 DCQLSI---LILLSCSVLDSF-------GELIPQ-PSNE  29
       |:||     ||||||||| |:|            |||
Db 116 DCELSSSGELLRSCSVKESLDLNESSSNPLVPDWESNE 154
```

RESULT   8
T28131
hypothetical protein ZK970.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28131
R;Berks, M.;
submitted to the EMBL Data Library, April 1995
A;Reference number: 220473
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: T28131
A;Molecule type: DNA
A;Residues: 1-147 <WIL>
A;Cross-references: EMBL:Z49073; PIDN:CAA88891.1; GSPDB:GN00020; CESP:ZK970.7
A;Experimental source: clone ZK970
C;Genetics:
A;Gene: ZK970.7
A;Map position: 2
A;Introns: 103/3

Query Match   33.8%; Score 49; DB 2; Length 147;
Best Local Similarity 37.5%; Pred. No. 4.6;
Matches   9; Conservative   7; Mismatches   8; Indels   0; Gaps   0;

```
Qy   3 CQLSILLLSCSVLDSFGELIPQP 26
       |  |   | |   |     |
Db   2 CRLAILALFSVAMVSVYGQDPPP 25
```

RESULT   9
T20204
hypothetical protein C54C8.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20204
R;Dobson, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: 219237
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Accession: T20204
A;Molecule type: DNA
A;Residues: 1-399 <WIL>
A;Cross-references: EMBL:Z83102; PIDN:CAB05466.1; GSPDB:GN00019; CESP:C54C8.4
A;Experimental source: clone C54C8
C;Genetics:
A;Gene: CESP:C54C8.4
A;Map position: 1

A;Introns: 87/1; 115/3; 189/1; 210/3; 266/3; 331/2

Query Match   32.0%; Score 47; DB 2; Length 399;
Best Local Similarity 62.5%; Pred. No. 25;
Matches   10; Conservative   3; Mismatches   3; Indels   0; Gaps   0;

```
Qy   3 CQLSILLLSCSVLDS 18
       | |: ||||||| |:
Db   6 CKLTIKLLLSCFVIYS 21
```

RESULT 10
F69063
tryptophan synthase, beta subunit homolog - Methanobacterium thermoautotrophicum (str
C;Species: Methanobacterium thermoautotrophicum
C;Accession: F69063
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
;Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A;Reference number: A69000; MUID:98037514
A;Accession: F69063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <MTH>
A;Cross-references: GB:AE000908; GB:AE000666; NID:g2622579; PIDN:AAB85951.1; PID:g262
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1476
A;Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homolog
C;Superfamily: complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
F;30-422/Domain: tryptophan synthase beta chain homology <TRPB>

Query Match   32.0%; Score 47; DB 2; Length 429;
Best Local Similarity 28.6%; Pred. No. 27;
Matches   14; Conservative   5; Mismatches   6; Indels 24; Gaps   2;

```
Qy   4 QLSILLLLSCSVLD-----------SF------------GELIPQPSN  28
       | | ||  ||:|              |            ||||  ||
Db 137 QWGTALSLACSLMDLQCKVYMVRVSFNQKPFRKTIMQLYGGEVVPPSN 185
```

RESULT 11
JE0378
DNA (cytosine-5-)methyltransferase (EC 2.1.1.37) - rat
DNA (cytosine-5-)-methyltransferase (Norway rat)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: JE0378
R;Kimura, H.; Takeda, T.; Tanaka, S.; Ogawa, T.; Shiota, K.
Biochem. Biophys. Res. Commun. 253, 495-501, 1998
A;Title: Expression of rat DNA (cytosine-5) methyltransferase (DNA MTase) in rodent t
A;Reference number: JE0378
A;Accession: JE0378
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1622 <KIM>
A;Cross-references: DDBJ:AB012214
A;Experimental source:
C;Keywords: methyltransferase

Query Match   32.0%; Score 47; DB 2; Length 1622;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches   8; Conservative   8; Mismatches   8; Indels   0; Gaps   0;

```
Qy    2 DCQLSILLLSCSVLDSFGELIPQ 25
        ||  ||||  |:: |   | :||
Db 1195 DCNVLLKIVMAGEVTNSLGQRLPQ 1218
```

RESULT 12

S23709
Interferon alpha-1 precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 22-Nov-1993
C;Accession: S23709, I46589
R;Mege, D.; Lefevre, F.; Labonnardiere, C.
J;Interferon Res. 11, 341-350, 1991
A;Title: The porcine family of interferon-omega: cloning, structural analysis, and funct
A;Reference number: S23709; MUID:92193689
A;Accession: S23709
A;Molecule type: mRNA
A;Residues: 1-189 <MEG>
A;Cross-references: EMBL:X57191; NID:g2062; PIDN:CAA40477.1; PID:g2063
R;Lefevre, F.; La Bonnardiere, C.
J;Interferon Res. 6, 349-360, 1986
A;Title: Molecular cloning and sequencing of a gene encoding biologically active porcine
A;Reference number: I46589; MUID:87035119
A;Accession: I46589
A;Superfamily: DNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-189 <LEF>
A;Cross-references: GB:M28623; NID:g164498; PIDN:AAA31053.1; PID:g164499
C;Genetics:
A;Gene: PoIFN-alpha-1
C;Superfamily: Interferon alpha

```
Qy
Best Local Similarity  31.3%;        Score 46;  DB 2;  Length 189;
Matches  10;  Conservative  5;  Mismatches  6;  Indels  0;  Gaps  0;

Qy   5  LSILILLLSCSVLDSFGELIPQ 25
        |:||:|||||||||||||||||
Db   8  LTALVLLSCNAICSLGCDLPQ 28
```

RESULT 13
T38260
hypothetical protein SPAC23C4.05c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999
C;Accession: T38260
R;Harris, D.; Squares, R.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21782
A;Accession: T38260
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-431 <HAR>
A;Experimental source: strain 972h-; cosmid c23C4
A;Cross-references: EMBL:Z99753; PIDN:CAB16876.1; GSPDB:GN00066; SPDB:SPAC23C4.05c
C;Genetics:
A;Gene: SPDB:SPAC23C4.05c
A;Map position: 1
A;Introns: 23/1

```
Qy
Best Local Similarity  72.7%;        Score 46;  DB 2;  Length 431;
Matches  8;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy  16  LDSFGELIPQP 26
        ||||||:|:|:|
Db  375 LESFGELVPVP 385
```

RESULT 14
B23595
DNA nucleotidylexotransferase (EC 2.7.7.31) - mouse
N;Alternate names: terminal addition enzyme; terminal deoxyribonucleotidyltransferase (1
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1999

C;Accession: B23595, S30235
R;Koiwai, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K.
Nucleic Acids Res. 14, 5777-5792, 1986
A;Title: Isolation and characterization of bovine and mouse terminal deoxynucleotidyl
A;Reference number: A93633; MUID:86286588
A;Accession: B23595
A;Molecule type: mRNA
A;Residues: 1-529 <KOI>
A;Cross-references: GB:X04123; NID:g54765; PIDN:CAA27735.1; PID:g54766
R;Doyen, N.; d'Andon, M.F.; Bentolila, L.A.; Nguyen, Q.T.; Rougeon, F.
Nucleic Acids Res. 21, 1187-1191, 1993
A;Title: Differential splicing in mouse thymus generates two forms of terminal deoxyn
A;Reference number: S30235; MUID:93219079
A;Accession: S30235
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-25,'M',27-98,'F',100-192,'R',194-286,'O',288-308,'E',310-366,'D',368-44
A;Cross-references: EMBL:X68670; NID:g287808; PIDN:CAA48634.1; PID:g287809
C;Superfamily: DNA nucleotidylexotransferase
C;Keywords: alternative splicing; magnesium; nucleotidyltransferase; nucleus

```
Qy
Best Local Similarity  31.3%;        Score 46;  DB 2;  Length 529;
Matches  10;  Conservative  3;  Mismatches  8;  Indels  0;  Gaps  0;

Qy   9  LLLSCSVLDSFGELIPQPSNE 29
        ||||:|:||:|| :|: :|
Db  374 LLLYCDILESTEFKFKQPSRK 394
```

RESULT 15
T27133
hypothetical protein Y53C12B.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27133
R;Kershaw, J.; Lennard, N.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z20316
A;Accession: T27133
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-793 <WIL>
A;Cross-references: EMBL:Z99278; PIDN:CAB16490.1; GSPDB:GN00020; CESP:Y53C12B.1
C;Genetics:
A;Gene: CESP:Y53C12B.1
A;Map position: 2
A;Introns: 51/2; 320/3; 464/2; 686/3

```
Qy
Best Local Similarity  34.0%;        Score 46;  DB 2;  Length 793;
Matches  16;  Conservative  5;  Mismatches  69;  Indels  18;  Gaps  3;

Qy   1  MDCQL------SIL---------LLLSCSVLDS--FGELIPQPSNE 29
        :|||      |:|           ||| |||||  |||||||  |:
Db  351 LDCQLIRGHTESVLSVVSPTWDTSLLASCSKDNSIIFWRLVHSPEND 397
```

Search completed: May 9, 2000, 22:27:18
Job time: 4157 sec

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 22:25:59 ; Search time 44.2 Seconds
9.490 Million cell updates/sec
(without alignments)

Title: US-09-104-340-3
Perfect score: 147
Sequence: 1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 143561 seqs, 14463640 residues

Total number of hits satisfying chosen parameters: 143561

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	100.0	983	1	US-08-167-919A-10	Sequence 10, Appl
2	147	100.0	983	2	US-08-449-645A-21	Sequence 21, Appl
3	147	100.0	983	2	US-08-702-367A-21	Sequence 21, Appl
4	147	100.0	983	3	US-08-715-106-10	Sequence 10, Appl
5	147	100.0	982	2	PCT-US95-04681-21	Sequence 21, Appl
6	51	34.7	983	2	US-08-673-789-4	Sequence 4, Appl
7	51	34.7	982	1	US-08-162-809-16	Sequence 16, Appl
8	47	32.0	983	1	US-08-167-919A-2	Sequence 2, Appl
9	47	32.0	14	3	US-08-715-106-2	Sequence 2, Appl
10	47	32.0	19	1	US-08-167-919A-3	Sequence 3, Appl
11	47	32.0	19	3	US-08-715-106-3	Sequence 3, Appl
12	45	30.6	128	4	US-08-476-275-4	Sequence 4, Appl
13	44	29.9	129	4	PCT-US95-07372-12	Sequence 12, Appl
14	44	29.9	316	2	US-08-827-291A-2	Sequence 2, Appl
15	43.5	29.6	448	1	US-08-207-904-17	Sequence 17, Appl
16	43.5	29.6	448	1	US-08-207-904-2	Sequence 2, Appl
17	43	29.3	235	4	PCT-US95-07302-2	Sequence 2, Appl
18	43	29.3	235	2	US-08-303-569B-5	Sequence 5, Appl
19	43	29.3	235	2	US-08-116-247-5	Sequence 5, Appl
20	43	29.3	515	2	US-08-468-249A-18	Sequence 18, Appl
21	43	29.3	585	1	US-08-142-439A-6	Sequence 6, Appl
22	43	29.3	585	2	US-08-142-551B-125	Sequence 125, App
23	43	29.3	585	2	US-08-869-477-6	Sequence 6, Appl
24	43	29.3	585	2	US-08-468-249A-19	Sequence 19, Appl
25	43	29.3	585	2	US-08-822-445-12	Sequence 12, Appl
26	43	29.3	3801	2	US-08-822-445-10	Sequence 10, Appl
27	42	28.6	3672	2	US-07-634-278-31	Sequence 31, Appl
28	42	28.6	128	1	US-07-946-421-26	Sequence 26, Appl
29	42	28.6	128	1	US-08-477-728-31	Sequence 31, Appl
30	42	28.6	128	1	US-08-474-040-31	Sequence 31, Appl
31	42	28.6	128	3	US-08-487-200-31	Sequence 31, Appl
32	42	28.6	128	2	US-08-444-644-15	Sequence 15, Appl
33	42	28.6	128	1	US-08-444-644-25	Sequence 25, Appl
34	41	27.9	277	1	US-08-118-270-68	Sequence 68, Appl
35	41	27.9	277	4	PCT-US93-08528-68	Sequence 68, Appl
36	41	27.9	339	1	US-08-396-957A-4	Sequence 4, Appl
37	41	27.9	687	2	US-08-449-645A-29	Sequence 29, Appl
38	41	27.9	687	2	US-08-702-367A-29	Sequence 29, Appl
39	41	27.9	915	1	US-08-452-734A-2	Sequence 2, Appl
40	41	27.9	915	4	PCT-US94-14989-2	Sequence 2, Appl
41	41	27.9	1242	2	US-08-580-326-33	Sequence 33, Appl
42	40.5	27.6	40	2	US-08-194-981E-14	Sequence 14, Appl
43	40.5	27.6	40	2	US-08-729-103-1	Sequence 1, Appl
44	40.5	27.6	158	2	US-08-453-862-2	Sequence 2, Appl
45	40.5	27.6	158	2	US-08-468-413-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-167-919A-10
Sequence 10, Application US/08167919A
Patent No. 5674691
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,919A
FILING DATE: 18-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-167-919A-10

Qy 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
      |||||||||||||||||||||||||||||
Db 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29

Query Match   100.0%; Score 147; DB 1; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
US-08-449-645A-21
; Sequence 21, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
;   APPLICANT: Fox, Gary M.
;   TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
;   TITLE OF INVENTION: Kinases
;   NUMBER OF SEQUENCES: 43
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Amgen Patent Operations/RBW
;   STREET: 1840 Dehavilland Drive
;   CITY: Thousand Oaks
;   STATE: California
;   COUNTRY: USA
;   ZIP: 91320
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/449,645A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Winter, Robert B.
;   REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 21:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 983 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-449-645A-21

Qy 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
      |||||||||||||||||||||||||||||
Db 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29

Query Match   100.0%; Score 147; DB 2; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
US-08-702-367A-21
; Sequence 21, Application US/0870236TA
; Patent No. 5981246
; GENERAL INFORMATION:
;   APPLICANT: Fox, Gary M.
;   TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
;   TITLE OF INVENTION: Kinases
;   NUMBER OF SEQUENCES: 43
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Amgen Patent Operations/RBW
;   STREET: 1840 Dehavilland Drive
;   CITY: Thousand Oaks
;   STATE: California

;   COUNTRY: USA
;   ZIP: 91320
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/702,367A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Winter, Robert B.
;   REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 21:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 983 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-702-367A-21

Qy 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
      |||||||||||||||||||||||||||||
Db 1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29

Query Match   100.0%; Score 147; DB 2; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
US-08-715-106-10
; Sequence 10, Application US/08715106
; Patent No. 6020306
; GENERAL INFORMATION:
;   APPLICANT: Boyd, Andrew W.
;   APPLICANT: Simpson, Richard J.
;   APPLICANT: Wicks, Ian
;   APPLICANT: Ward, Larry D.
;   APPLICANT: Wilkinson, David
;   TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
;   TITLE OF INVENTION: AND USE THEREOF
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Scully, Scott, Murphy & Presser
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/715,106
;   FILING DATE: 18-SEP-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/167,919
;   FILING DATE: 18-APR-1994
;   APPLICATION NUMBER: PK6841 (AU)
;   FILING DATE: 21-JUN-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PK9992 (AU)
;   FILING DATE: 12-DEC-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/AU92/00294
;   FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:
  NAME: DiGiglio, Frank S.
  REGISTRATION NUMBER: 31,346
  REFERENCE/DOCKET NUMBER: 9159
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (516) 742-4343
  TELEFAX: (516) 742-4366
  TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 983 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-715-106-10

```
Qy   1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
       |||||||||||||||||||||||||||||
Db   1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
```

Query Match   100.0%; Score 147; DB 3; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches   29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
PCT-US95-04681-21
Sequence 21, Application PC/TUS9504681
GENERAL INFORMATION:
  APPLICANT: Fox, Gary M.
  TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
  TITLE OF INVENTION: Kinases
  NUMBER OF SEQUENCES: 28
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Amgen Patent Operations/RBW
  STREET: 1840 Dehavilland Drive
  CITY: Thousand Oaks
  STATE: California
  COUNTRY: USA
  ZIP: 91320
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/04681
  FILING DATE:
  CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
  NAME: Winter, Robert B.
  REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 983 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
PCT-US95-04681-21

```
Qy   1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
       |||||||||||||||||||||||||||||
Db   1 MDCQLSILLLLSCSVLDSFGELIPQPSNE 29
```

Query Match   100.0%; Score 147; DB 4; Length 983;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches   29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6

US-08-673-789-4
Sequence 4, Application US/08673789
Patent NO. 5814479
GENERAL INFORMATION:
  APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS;
  APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
  APPLICANT: GEORGE, F.
  TITLE OF INVENTION: BSK RECEPTOR LIKE
  TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
  TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
  TITLE OF INVENTION: METHODS
  NUMBER OF SEQUENCES: 14
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: MORGAN & FINNEGAN
  STREET: 345 PARK AVENUE
  CITY: NEW YORK
  STATE: NEW YORK
  COUNTRY: USA
  ZIP: 10154
  COMPUTER READABLE FORM:
  MEDIUM TYPE: FLOPPY DISK
  COMPUTER: IBM PC COMPATIBLE
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: WORDPERFECT 5.1
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/673,789
  FILING DATE:
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/177,812
  FILING DATE: 04-JAN-1994
  ATTORNEY/AGENT INFORMATION:
  NAME: CAROL M. GRUPPI
  REGISTRATION NUMBER: 37,341
  REFERENCE/DOCKET NUMBER: 2026-4105
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (212) 758-4800
  TELEFAX: (212) 751-6849
  TELEX: 421792
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 982
  TYPE: AMINO ACID
  STRANDEDNESS: UNKNOWN
  TOPOLOGY: UNKNOWN
US-08-673-789-4

```
Qy   9 LILLSCSVLDSFGELIPQPSNE 29
       ||| ||:|||||||| :|| ||
Db   8 LILLCAALGSAGRLSARPGNE 28
```

Query Match   34.7%; Score 51; DB 2; Length 982;
Best Local Similarity 52.4%; Pred. No. 8.7;
Matches   11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

RESULT 7
US-08-162-809-16
Sequence 16, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
  APPLICANT: Pasquale, Elena B.
  APPLICANT: Sajjadi, Fereydoun G.
  TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
  NUMBER OF SEQUENCES: 26
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: CAMPBELL AND FLORES
  STREET: 4370 La Jolla Village Drive, Suite 700
  CITY: San Diego
  STATE: California
  COUNTRY: United States of America

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-162-809-16

Query Match
Best Local Similarity   34.7%;   Score 51;   DB 1;   Length 983;
Matches 11;   Conservative   52.4%;   Pred. No. 8.7;   Mismatches 8;   Indels 0;   Gaps 0;

```
Qy    9 LLLSCCSVLDSFGELIPQPSNE 29
          |||  |||||| || |||
Db    8 LLLCAALGSAGRLSARPGNE 28
```

RESULT 8
US-08-167-919A-2
Sequence 2, Application US/0816791 9A
Patent No. 5674691
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,919A
FILING DATE: 18-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992

ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-167-919A-2

Query Match
Best Local Similarity   32.0%;   Score 47;   DB 1;   Length 14;
Matches 9;   Conservative   100.0%;   Pred. No. 0.26;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy   21 ELIPQPSNE 29
          |||||||||
Db    1 ELIPQPSNE 9
```

RESULT 9
US-08-715-106-2
Sequence 2, Application US/08715106
Patent No. 6020306
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,106
FILING DATE: 18-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,919
FILING DATE: 18-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343

RESULT 10
US-08-167-919A-3
Sequence 3, Application US/0816791A
Patent No. 5674691
; GENERAL INFORMATION:
; APPLICANT: Boyd, Andrew W.
; APPLICANT: Simpson, Richard J.
; APPLICANT: Wicks, Ian
; APPLICANT: Ward, Larry D.
; APPLICANT: Wilkinson, David
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,919A
; FILING DATE: 18-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK6841 (AU)
; FILING DATE: 21-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK9992 (AU)
; FILING DATE: 12-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00294
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9159
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-715-106-2

Qy   21 ELIPQPSNE 29
        |||||||||
Db   1 ELIPQPSNE 9

Query Match   32.0%;
Best Local Similarity   100.0%; Score 47; DB 3; Length 14;
Matches   9; Conservative   0; Pred. No. 0.26; Mismatches   0; Indels   0; Gaps   0;

; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

RESULT 11
US-08-715-106-3
Sequence 3, Application US/08715106
Patent No. 6020306
; GENERAL INFORMATION:
; APPLICANT: Boyd, Andrew W.
; APPLICANT: Simpson, Richard J.
; APPLICANT: Wicks, Ian
; APPLICANT: Ward, Larry D.
; APPLICANT: Wilkinson, David
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,106
; FILING DATE: 18-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,919
; FILING DATE: 18-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK6841 (AU)
; FILING DATE: 21-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK9992 (AU)
; FILING DATE: 12-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00294
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9159
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-715-106-3

Qy   21 ELIPQPSNE 29
        |||||||||
Db   1 ELIPQPSNE.9

Query Match   32.0%;
Best Local Similarity   100.0%; Score 47; DB 1; Length 19;
Matches   9; Conservative   0; Pred. No. 0.37; Mismatches   0; Indels   0; Gaps   0;

; MOLECULE TYPE: protein
US-08-167-919A-3

Qy 21 ELIPQPSNE 29
        |||||||||
Db   1 ELIPQPSNE 9

Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 12
US-08-476-275-4
Sequence 4, Application US/08476275
Patent No. 5776456
GENERAL INFORMATION:
  APPLICANT: Anderson, Darrell R.
  APPLICANT: Hanna, Nabil
  APPLICANT: Leonard, John E.
  APPLICANT: Newman, Roland A.
  APPLICANT: Reff, Mitchell E.
  APPLICANT: Rastetter, William H.
  TITLE OF INVENTION: Therapeutic Application of Chimeric and
  TITLE OF INVENTION: Radiolabeled Antibodies to Human B Lymphocyte Restricted
  TITLE OF INVENTION: Differentiation Antigen for the Treatment of B-Cell
  TITLE OF INVENTION: Lymphoma
  NUMBER OF SEQUENCES: 11
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
    STREET: 699 Prince St.
    CITY: Alexandria
    STATE: VA
    COUNTRY: USA
    ZIP: 22314
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/476,275
    FILING DATE: 07-JUN-1995
    CLASSIFICATION: 424
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/149,099
    FILING DATE: 03-NOV-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/978,891
    FILING DATE: 13-NOV-1992
  ATTORNEY/AGENT INFORMATION:
    NAME: Teskin, Robin L.
    REGISTRATION NUMBER: 35,030
    REFERENCE/DOCKET NUMBER: 012712-155
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 703-836-6620
    TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 128 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein

Query Match   30.6%; Score 45; DB 1; Length 128;
Best Local Similarity 44.0%; Pred. No. 6.7;
Matches 11; Conservative 7; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MDCQLSIL--LLLSCSVLDSFGELI 23
       ||:|:| |:|::|:|:::|
Db 1 MDFQVQIFSFLLISASVIMSRGQIV 25

RESULT 13
PCT-US95-07372-12
; Sequence 12, Application PC/TUS9507372

GENERAL INFORMATION:
  APPLICANT: Oklahoma Medical Research Foundation
  TITLE OF INVENTION: Calcium Binding Recombinant
  TITLE OF INVENTION: Antibody Against Protein C
  NUMBER OF SEQUENCES: 12
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Patrea L. Pabst
    STREET: 2800 One Atlantic Center
    STREET: 1201 West Peachtree Street
    CITY: Atlanta
    STATE: Georgia
    COUNTRY: USA
    ZIP: 30309-3450
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: PCT/US95/07372
    FILING DATE:
    CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Pabst, Patrea L.
    REGISTRATION NUMBER: 31,284
    REFERENCE/DOCKET NUMBER: OMRF106CIP
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (404) 873-8794
    TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 12:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 129 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: peptide
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  FRAGMENT TYPE: N-terminal
  ORIGINAL SOURCE:
    ORGANISM: HPC-4 Light Chain Variable Region (VL Kappa)
  FEATURE:
    NAME/KEY: misc-feature
    LOCATION: 23..129
    OTHER INFORMATION: /note= "Gln at position 23 starts
    OTHER INFORMATION: mature peptide."
PCT-US95-07372-12

Query Match   29.9%; Score 44; DB 4; Length 129;
Best Local Similarity 48.0%; Pred. No. 9.7;
Matches 12; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

Qy 1 MDCQLSI--LLLLSCSVLDSFGELI 23
       |||:| |:|::|:|:::|
Db 1 MDFQVQIFSFLLISASVIMSRGQIV 25

RESULT 14
US-08-827-291A-2
; Sequence 2, Application US/08272291A
; Patent No. 5874243
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Sathe, Ganesh
;   TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY:

```
;   ZIP:  19406
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/827,291A
;       FILING DATE:  28-MAR-1997
;       CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;       NAME:  King, William T
;       REGISTRATION NUMBER:  30,954
;       REFERENCE/DOCKET NUMBER:  GP50001
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610-270-5015
;       TELEFAX:  610-270-5090
;       TELEX:
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;       LENGTH:  316 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
;   US-08-827-291A-2
```

```
Qy       3  COLSILILLSCSVLDSFGELI  23
            |:| ||:|||:| :|
Db     179  CELPSLILSCNDTSIFEKVI  199
```

```
Query Match             29.9%;  Score 44;  DB 2;  Length 316;
Best Local Similarity   42.9%;  Pred. No. 27;
Matches    9;  Conservative   5;  Mismatches   7;  Indels   0;  Gaps   0;
```

```
RESULT 15
US-08-207-904-2
Sequence 2, Application US/0820790
Patent No. 5477002
GENERAL INFORMATION:
;   APPLICANT:  Tuttle, AnnMarie
;   APPLICANT:  Crossland, Lyle D.
;   TITLE OF INVENTION:  Anther-Specific cDNA Sequences, Genomic
;   TITLE OF INVENTION:  DNA Sequences and Recombinant DNA Sequences
;   NUMBER OF SEQUENCES:  21
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CIBA-GEIGY Corporation
;       STREET:  7 Skyline Drive
;       CITY:  Hawthorne
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10532
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/207,904
;       FILING DATE:
;       CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/908,242
;       FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;       NAME:  Lazar, Steven R.
;       REGISTRATION NUMBER:  32,618
;       REFERENCE/DOCKET NUMBER:  CGC 1624
```

```
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (919)541-8615
;       TELEFAX:  (919)541-8689
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;       LENGTH:  448 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
;   US-08-207-904-2
```

```
Qy       1  MDCQLSILILLSCSV--LDSFGELIPQP  26
            :|| | :|:|:   || |: | |
Db      87  LDCNASGIVLMRAETEAKILDDLGDFSPSP  115
```

```
Query Match             29.6%;  Score 43.5;  DB 1;  Length 448;
Best Local Similarity   31.0%;  Pred. No. 48;
Matches    9;  Conservative   5;  Mismatches  12;  Indels   3;  Gaps   1;
```

```
Search completed:  May  9, 2000, 22:26:00
Job time:  4149 sec
```

This Page Blank (uspto)

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.
GenCore version 4.5

Run on: May 9, 2000, 21:52:03 ; Search time 50.71 Seconds
                                    (without alignments)
                                    13.546 Million cell updates/sec

Title:
Perfect score: 147
Sequence: US-09-104-340-3
                  147
                  1 MDCQLSILLLSCSVLDSFGEELIPQPSNE 29

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	100.0	983	1	R31466	HEK polypeptide. R
2	55	37.4	288	1	W88378	Human sperm-specif
3	55	37.4	323	1	W88380	Human sperm-specif
4	51	34.7	983	1	R75711	Eph-related PTK Ce
5	48.5	33.0	2458	1	R07640	Deduced protein se
6	48.5	33.0	2458	1	R04031	Full length T4 enc
7	47.5	32.3	242	1	W58379	Human immunomodula
8	46	31.3	189	1	W83902	Porcine interferon
9	46	31.3	189	1	W73231	Porcine interferon
10	45.5	31.0	316	1	W72321	Human cytochrome P
11	45.5	31.0	490	1	R81465	Human derived cyto
12	45.5	31.0	490	1	R93168	Human cytochrome P
13	45.5	30.6	114	1	W27850	Staphylococcus aur
14	45	30.6	128	1	R55214	Murine variable re
15	45	30.6	908	1	W41568	Human metabotropic
16	44	29.9	129	1	R88108	Murine anti-Protei
17	44	29.9	316	1	W75960	Human olfactory OL
18	43.5	29.6	448	1	R47475	Polypeptide encode
19	43.5	29.6	448	1	R46929	Tobacco Ant32 geno
20	43	29.3	128	1	R90690	Mouse 5C7.29 monoc
21	43	29.3	129	1	P82937	Variable region of
22	43	29.3	129	1	P70626	Sequence encoded b
23	43	29.3	129	1	W16342	Variable region of
24	43	29.3	129	1	W10241	L6 antibody V-kapp
25	43	29.3	129	1	W10587	V region of the L6
26	43	29.3	129	1	W47512	Mouse L6 antibody
27	43	29.3	129	1	W41069	Mouse L6 antibody
28	43	29.3	129	1	W47519	Mouse L6 antibody
29	43	29.3	129	1	W89537	Anti-cancer antibo
30	43	29.3	135	1	R08347	Light chain variab
31	43	29.3	135	1	R13060	Monoclonal antibod
32	43	29.3	235	1	W41410	Humanised light ch
33	43	29.3	515	1	R27704	Opossum kidney PTH
34	43	29.3	515	1	R92275	Opossum kidney PTH
35	43	29.3	515	1	W73314	Parathyroid hormon
36	43	29.3	565	1	R76700	EIET enterotoxin e
37	43	29.3	565	1	R76699	EIET enterotoxin e
38	43	29.3	585	1	R27705	ShET2 enterotoxin
39	43	29.3	585	1	R92276	Opossum kidney PTH
40	43	29.3	585	1	W73315	Opossum kidney PTH
41	43	29.3	611	1	W83428	Parathyroid hormon
42	43	29.3	1763	1	W83431	Munc13-1-interacti
43	43	29.3	2001	1	W23596	Rat Munc13-1, Scor
44	43	29.3	3672	1	W31950	Human LYST1 longer
45	43	29.3	3801	1	W31949	Human bg protein a

ALIGNMENTS

RESULT 1

ID   R31466 standard; Protein; 983 AA.
AC   R31466;
DT   24-MAY-1993   (first entry)
DE   HEK polypeptide.
KW   HEK polypeptide. R
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   peptide   1..20
FT   /note= "Signal peptide"
FT   protein   21..39
FT   /note= "Purified HEK protein #1"
FT   protein   840..860
FT   /note= "Purified HEK protein #2"
FT   modified_site   232..234
FT   /note= "N-link glycosylation"
FT   modified_site   337..339
FT   /note= "N-link glycosylation"
FT   modified_site   391..393
FT   /note= "N-link glycosylation"
FT   modified_site   404..406
FT   /note= "N-link glycosylation"
FT   modified_site   493..495
FT   /note= "N-link glycosylation"
FT   domain   542..565
FT   /note= "Transmembrane domain"
FT   binding_site   628
FT   /note= "ATP binding site"
FT   binding_site   630
FT   /note= "ATP binding site"
FT   binding_site   633
FT   /note= "ATP binding site"
FT   modified_site   779
FT   /note= "Putative autophosphorylation site"
PN   WO9300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU0294.
PR   19-JUN-1992; AU-006841.
PR   21-JUN-1991; AU-009992.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   WPI; 93-036373/04.
DR   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence represets human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the

CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence 983 AA;

Qy   Query Match   100.0%;   Score 147;   DB 1;   Length 983;
     Best Local Similarity   100.0%;   Pred. No. 6.2e-14;
     Matches   29;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29
        |||||||||||||||||||||||||||||
Db   1 MDCQLSILILLSCSVLDSFGELIPQPSNE 29

RESULT 2
ID   W88378 standard; Protein; 288 AA.
AC   W88378;
DT   26-APR-1999   (first entry)
DE   Human sperm-specific antigen 4.1 (partial sequence).
KW   Sperm-specific antigen 4.1; vaccine; contraceptive; human.
OS   Homo sapiens.
PN   WO9858954-A2.
PD   30-DEC-1998;
PF   17-JUN-1998; E04016.
PR   19-JUN-1997; EP-201863.
PA   (ALKU) AKZO NOBEL NV.
PI   Aitken RJ, Grootenhuis AJ, Van Duin M;
DR   WPI; 99-095361/08.
DR   N-PSDB; X06777.
PT   New sperm proteins 4.1 and 6.7 - useful as a contraceptive vaccine
PS   Claim 8; Page 28-29; 44pp; English.
CC   This is a partial sequence of human sperm specific antigen 4.1.
CC   The sequence was deduced from a cDNA clone (see X06777) isolated
CC   from a human testicular cDNA library using marmoset 4.1 cDNA
CC   as probe. A full-length sequence for protein 4.1 is provided in
CC   W88380. Human and marmoset sperm proteins 4.1 and 6.7 (see
CC   W88373-80) and nucleotide sequences (see X06772-79) encoding them
CC   are new. The proteins are sperm specific and can be used in a
CC   contraceptive vaccine to prevent or reduce fertility in a subject
CC   without the autoimmune reactions that are a serious side-effect of
CC   prior art female contraceptive vaccines. Expression vectors and
CC   host cells used in production of recombinant sperm proteins are
CC   provided.
SQ   Sequence 288 AA;

Qy   Query Match   37.4%;   Score 55;   DB 1;   Length 288;
     Best Local Similarity   58.3%;   Pred. No. 0.88;
     Matches   14;   Conservative   3;   Mismatches   5;   Indels   2;   Gaps   1;

Qy   6 SILILLS--CSVLDSFGELIPQPS 27
        |:||| | ||:| ||||||| |||
Db   26 SVLLALPAWCRVLETFLELIQOPS 49

RESULT 3
W88380
ID   W88380 standard; Protein; 323 AA.
AC   W88380;
DT   26-APR-1999   (first entry)
DE   Human sperm-specific antigen 4.1.
KW   Sperm-specific antigen 4.1; vaccine; contraceptive; human.
OS   Homo sapiens.
PN   WO9858954-A2.
PD   30-DEC-1998;
PF   17-JUN-1998; E04016.
PR   19-JUN-1997; EP-201863.
PA   (ALKU) AKZO NOBEL NV.
PI   Aitken RJ, Grootenhuis AJ, Van Duin M;
DR   WPI; 99-095361/08.
DR   N-PSDB; X06779.
PT   New sperm proteins 4.1 and 6.7 - useful as a contraceptive vaccine
PS   Claim 8; Page 32-33; 44pp; English.
CC   This is a full-length sequence of human sperm specific antigen 4.1.
CC   The sequence was deduced from a cDNA clone (see X06777) isolated
CC   from a human testicular cDNA library using marmoset 4.1 cDNA
CC   as probe. Human and marmoset sperm proteins 4.1 and 6.7 (see
CC   W88373-80) and nucleotide sequences (see X06772-79) encoding them
CC   are new. The proteins are sperm specific and can be used in a
CC   contraceptive vaccine to prevent or reduce fertility in a subject
CC   without the autoimmune reactions that are a serious side-effect of
CC   prior art female contraceptive vaccines. Expression vectors and
CC   host cells used in production of recombinant sperm proteins are
CC   provided.
SQ   Sequence 323 AA;

Qy   Query Match   37.4%;   Score 55;   DB 1;   Length 323;
     Best Local Similarity   58.3%;   Pred. No. 1;
     Matches   14;   Conservative   3;   Mismatches   5;   Indels   2;   Gaps   1;

Qy   6 SILILLS--CSVLDSFGELIPQPS 27
        |:||| | ||:| ||||||| |||
Db   61 SVLLALPAWCRVLETFLELIQOPS 84

RESULT 4
R75711
ID   R75711 standard; Protein; 983 AA.
AC   R75711;
DT   11-NOV-1995
DE   Cek4; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   Eph-related PTK Cek4.
KW   prognosis.
OS   Gallus sp.
PN   WO9515375-A.
PD   08-JUN-1995.
PR   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;
DR   WPI; 95-215256/28.
DR   N-PSDB; Q90659.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Disclosure; Page 85-89; 129pp; English.
CC   Probes derived from the EPH-related PTKs Cek4 (Q90659), and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and
CC   embryonic tissues and also in the adult brain and retina.
SQ   Sequence 983 AA;

Qy   Query Match   34.7%;   Score 51;   DB 1;   Length 983;
     Best Local Similarity   52.4%;   Pred. No. 14;
     Matches   11;   Conservative   2;   Mismatches   8;   Indels   0;   Gaps   0;

Qy   6 SILILLS--CSVLDSFGELIPQPS 27
        |:||| | :| :| |:|| | |||
Db   61 SVLLALPAWCRVLETFLELIQOPS 84

RESULT 5
R07640
ID   R07640 standard; protein; 2458 AA.
AC   R07640;
DT   20-DEC-1990   (first entry)
DE   Deduced protein sequence of p170-2 comprising T4.
KW   plasmid p170-2; soluble T4 protein; AIDS; ARC; HIV.
FH   Key   Location/Qualifiers
FT   protein   400..858
FT   /label=T4 surface glycoprotein
PN   WO9008198-A.

Qy   9 LLLSCSVLDSFGELIPQPSNE 29
        ||| :|||:||||:|| :| ||
Db   8 LLLLCAALGSAGRLSARPGNE 28

PD   26-JUL-1990.
PD   18-JAN-1990; U00358.
PF   18-JAN-1989; US-300096.
PA   (HARD) HARVARD COLLEGE.
PI   Letvin NA;
DR   WPI; 90-25404O/33.
PN   N-PSDB; Q05607.
PT   Treating or preventing AIDS, ARC or HIV infection - by
PT   administering an immunologically effective amt. of soluble T4
CC   protein.
PS   Disclosure; Fig 1; 121pp; English.
CC   Entire sequence from T4-encoding plasmid p170-2. It is almost
CC   identical to the sequence published by Madden et al.(1985) with the
CC   exception of three codon changes. At T4 amino acid residue 3,
CC   (posn.403 of entire sequence) Lys is encoded in stead of Asn. At
CC   posn. 64, (posn.464) Arg replaces Trp and at posn. 231, (posn. 631)
CC   Ser replaces Phe.
CC   Soluble T4 can be produced by truncating the CDS to remove the
CC   transmembrane and cytoplasmic domains. The soluble forms may be
CC   modified to increase their immunogenicity by addition of an adjuvant
CC   such as incomplete Freund's adjuvant. The T4 interferes with HIV/T4
CC   interaction and elicits anti-soluble T4 antibody production.
CC   See also Q05608.
SQ   Sequence 2458 AA;

Query Match   33.0%; Score 48.5; DB 1; Length 2458;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy   4 QLSILLLSCSVLDSFGELI-PQPSN 28
         |:||||| |:| |: |||:
Db 2268 QHCIILLLSCHPXDAFLXLVSTQPSH 2293

RESULT 6
ID   R04031 standard; protein; 2458 AA.
AC   R04031;
DT   29-MAY-1990 (first entry)
DE   Full length T4 encoded by plasmid p170-2.
KW   Soluble T4; p170-2; anti-retroviral agent; AIDS; ARC; HIV; AZT.
FH   Key   Location/Qualifiers
FT   misc_difference 425
FT   /note="Asp of Madden et al replaced by Lys"
FT   misc_difference 423
FT   /note="Trp of Madden et al replaced by Arg"
FT   misc_difference 653
FT   /note="Phe of Madden et al replaced by Ser"
DR   WPI; 90-007302/01.
DR   WO8911860-A.
PN   N-PSDB; Q03005.
PD   14-DEC-1989.
PD   08-JUN-1989; U02453.
PF   10-JUN-1988; US-204645.
PR   20-APR-1989; US-341080.
PA   (GEHO-) General Hospital Corp   (BIOJ) Biogen Inc.
PI   Fisher RA, Schooley RT, Hirsch MS, Johnson VA, Walker BD;
PT   Combinations of soluble T4 protein and anti-retroviral agent -
PT   having synergistic activity in treatment and prevention of AIDS,
PT   ARC and HIV infection.
PS   Disclosure; fig 1; 100pp; English.
CC   The sequence differs from that determined by Madden et al., [Cell, 42
CC   pp. 93-104 (1985)] in three places due to three nucleotide substitutions.
CC   The Asp reported at position 3 by Madden et al. was the result of a
CC   sequencing error [DR Littman et al. Cell, 55, p.541 (1988)].
CC   X = stop codon.
CC   The sequence was deduced from the cDNA insert of p170-2.
CC   Soluble T4 constructs may be produced by truncating this sequence to give
CC   fragments from position 400 to 799, removing the transmembrane and
CC   intracytoplasmic domains whilst retaining the extracellular region
CC   responsible for HIV binding. The sol. T4 is combined with an anti-viral
CC   agent such as AZT.

CC   See also Q03006.
SQ   Sequence 2458 AA;

Query Match   33.0%; Score 48.5; DB 1; Length 2458;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy   4 QLSILLLSCSVLDSFGELI-PQPSN 28
         |:||||| |:| |: |||:
Db 2268 QHCIILLLSCHPXDAFLXLVSTQPSH 2293

RESULT 7
ID   W58379 standard; protein; 242 AA.
AC   W58379;
DT   08-SEP-1998 (first entry)
DE   Human immunomodulatory protein BL-1 3.
KW   Human; BL-1; immunomodulatory protein; vaccine; immunotherapy; HIV;
KW   infection; cancer; autoimmune disease.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   Misc_difference 1. .242
FT   /note= "the protein is decoded from the third reading
FT   frame of BL-1 cDNA and so contains stop codons
FT   (i.e. all X's are stop codons"
DR   WPI; 98-261495/23.
DR   WO9817799-A1.
PN   N-PSDB; V30915.
PD   30-APR-1998.
PF   23-OCT-1997; U19502.
PR   23-OCT-1996; US-028613.
PA   (APOL-) APOLLON INC.
PA   (UYPE-) UNIV PENNSYLVANIA.
PI   Bagarazzi ML, Boyer JD, Kim JJ, Wang B, Weiner DB;
PT   New compositions for immuno-therapy and protection - comprise
PT   nucleotide sequences encoding an immuno-modulating protein and an
PT   antigen, used for e.g. infections, cancer or auto-immune diseases
PS   Example 6; Fig 14; 136pp; English.
CC   The present sequence represents the protein from the third reading
CC   frame of human immunomodulatory protein BL-1 encoding cDNA (given in
CC   V30915). The present invention describes a plasmid which comprises a
CC   nucleotide sequence (NS) that encodes:(a) an immunomodulatory protein
CC   selected from interleukin (IL)-12, granulocyte-macrophage colony
CC   stimulating factor (GM-CSF), IL-1, tumour necrosis factor (TNF)-alpha,
CC   TNF- beta, IL-2, IL-4, IL-5, IL-10, IL-15, IL-18 and BL-1 operably
CC   linked to regulatory elements. Products of the present invention can be
CC   used to induce an immune response to an antigen such as a pathogen
CC   antigen, a hyperproliferative disease-associated antigen, and antigen
CC   linked to cells associated with autoimmune diseases or an allergen.
CC   They can be used for immunotherapy or to provide a protective immune
CC   response. In particular, they can be used for treating subjects with
CC   an allergic reaction, pathogen infection, hyperproliferative disease
CC   such as cancer or psoriasis or autoimmune diseases e.g. rheumatoid
CC   arthritis, multiple sclerosis, Sjogren's syndrome, sarcoidosis, Insulin
CC   dependent diabetes mellitus, autoimmune thyroiditis, ankylosing
CC   spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis,
CC   vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative
CC   colitis, Grave's disease, autoimmune thrombocytopenia, asthma and
CC   pernicious anaemia.
SQ   Sequence 242 AA;

Query Match   32.3%; Score 47.5; DB 1; Length 242;
Best Local Similarity 50.0%; Pred. No. 9.6;
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy   4 QLSILLLSCSVLDSFGELI-PQPSN 28
        |:||||| |:| |: |||:
Db 149 QHCIILLLSCHPQDAFLXLVSTQPSH 174

```
RESULT 8
W83902

ID   W83902 standard; protein; 189 AA.
AC   W83902;
DT   15-FEB-1999 (first entry)
DE   Porcine interferon-alpha 1.
KW   Interferon-alpha 1; leukocyte interferon; virucide; antiviral;
KW   antitumour; pig.
OS   Sus scrofa.
FH   Key             Location/Qualifiers
FH
FT   Peptide         1..23
FT                   /label= Sig_peptide
FT   Protein         24..189
FT                   /label= Mat_protein
FT                   /note= "Claim 1"
PN   US5831023-A.
PD   03-NOV-1998.
PF   19-MAY-1995; 444454.
PR   02-OCT-1987; US-104461.
PR   08-MAR-1982; US-355298.
PR   01-NOV-1982; US-438128.
PR   23-AUG-1991; US-749371.
PR   21-SEP-1992; US-949327.
PR   19-MAY-1995; US-444454.
PI   Capon DJ, Goeddel DV;
PA   (GETH ) GENENTECH INC.
DR   WPI; 99-007993/01.
DR   N-PSDB; V68167.
PT   Animal interferon polypeptides - useful as antiviral and antitumour
PT   agents
PS   Claim 1; Fig 14A; 48pp; English.
CC   This is the amino acid sequence of porcine interferon alpha-1, as
CC   deduced from the open reading frame of a cDNA clone (see V68167).
CC   The invention provides bovine, murine, porcine, feline, laprine
CC   and rat mature alpha, beta and gamma interferons and their
CC   propeptides (see W83895-905, W83918 and W70582-85). Recombinant
CC   DNA methods can be utilised to prepare these non-human animal
CC   interferons in amounts sufficient to enable the determination of
CC   their biochemical properties and bioactivities, and hence to allow
CC   efficient production for commercial or biological exploitation.
CC   Interferons are useful as antiviral and antitumour agents.
SQ   Sequence 189 AA;

Qy   5 LSILLLLSCSVLDSFGELIPQ 25
       |: |:||||: :| | ||
Db   8 LTALVLLSCNAICSLGCDLPQ 28

Query Match        31.3%;  Score 46;  DB 1;  Length 189;
Best Local Similarity  47.6%;  Pred. No. 12;
Matches   10;  Conservative  5;  Mismatches  6;  Indels  0;  Gaps  0;
```

```
RESULT 9
W73231

ID   W73231 standard; protein; 189 AA.
AC   W73231;
DT   01-MAR-1999 (first entry)
DE   Porcine interferon-alpha 1; leukocyte interferon; virucide; antiviral;
KW   antitumour; pig.
OS   Sus scrofa.
FH   Key             Location/Qualifiers
FH
FT   Peptide         1..23
FT                   /label= Sig_peptide
FT   Protein         24..189
FT                   /label= Mat_protein
PN   US5827694-A.
PD   27-OCT-1998.
PF   19-MAY-1995; 446171.
PR   02-OCT-1987; US-104461.
PR   08-MAR-1982; US-395298.
PR   01-NOV-1982; US-438128.
```

```
PR   23-AUG-1991; US-749371.
PR   21-SEP-1992; US-949327.
PR   19-MAY-1995; US-446171.
PI   Capon DJ, Goeddel DV;
PA   (GETH ) GENENTECH INC.
DR   WPI; 98-593994/50.
DR   N-PSDB; V08184.
PT   DNA encoding non-human interferon polypeptide(s) - useful for
PT   producing recombinant polypeptide(s) to determine their
PT   bioactivity
PS   Claim 6; Fig 14A; 59pp; English.
CC   This is the amino acid sequence of porcine interferon alpha-1,
CC   deduced from a cDNA (see V08184). The invention provides
CC   bovine, porcine, feline and rabbit mature interferons and their
CC   propeptides (see W73224-35). Recombinant DNA methods can be used
CC   to prepare these non-human animal interferons in amounts sufficient
CC   to enable the determination of their biochemical properties and
CC   bioactivities, and hence to allow efficient production for
CC   commercial or biological exploitation. Interferons are useful as
CC   antiviral and antitumour agents. A replicable expression vehicle
CC   comprising claimed non-human animal interferon nucleic acid, a
CC   process for producing such an interferon, and a microorganism
CC   (especially Escherichia coli) or cell culture transfected with the
CC   nucleic acid are claimed.
SQ   Sequence 189 AA;

Qy   5 LSILLLLSCSVLDSFGELIPQ 25
       |: |:||||: :| | ||
Db   8 LTALVLLSCNAICSLGCDLPQ 28

Query Match        31.3%;  Score 46;  DB 1;  Length 189;
Best Local Similarity  47.6%;  Pred. No. 12;
Matches   10;  Conservative  5;  Mismatches  6;  Indels  0;  Gaps  0;
```

```
RESULT 10
R72361

ID   R72361 standard; protein; 490 AA.
AC   R72361;
DT   10-NOV-1995 (first entry)
DE   Human cytochrome P450 molecular species 2C9 protein.
KW   Human cytochrome P450; amplification; PCR; primer; expression vector;
KW   yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
KW   carcinogen; mutagen; liver metabolism.
OS   Homo sapiens.
PN   EP-644267-A.
PD   22-MAR-1995.
PF   20-JUL-1994; 111298.
PR   20-JUL-1993; JP-201120.
PR   21-JUL-1993; JP-180246.
PR   30-JUL-1993; JP-208279.
PI   (HAYA/) HAYASHI K.
PA   (SUMO ) SUMITOMO CHEM CO LTD.
PI   Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T;
PI   Yabusaki Y;
DR   WPI; 95-116991/16.
DR   N-PSDB; Q87715.
PT   Evaluation of safety of a chemical cpd. - using recombinant yeast
PT   expressing human cytochrome p450 and a yeast NADPH-P450 reductase
PS   Examples; Page 23-25; 124pp; English.
CC   The amino acid sequence of the human cytochrome P450 species 2C9.
CC   cDNA was amplified by PCR using the primers Q87735-8. The product was
CC   cloned into the yeast expression vectors pAAH5N or pAHRR to produce the
CC   vectors p2C9 for the expression of the cytochrome P450 alone or p2C9R
CC   for co-expression with the yeast NADPH-P450 reductase.
CC   The vectors are used in a method for evaluating the safety of a chemical
CC   compound by reacting the chemical compound with recombinantly produced
CC   human cytochrome P450 molecular species 1A2 (Q87714), 2C9, 2E1 (Q87716)
CC   or 3A4 (Q87717), or their auxiliary species and variants (Q87718-32), and
CC   yeast NADPH-P450 reductase, either as a fused protein or in cell
CC   extracts, and analysing the resulting metabolite to assess the safety of
CC   the chemical compound. The method is useful for determining whether the

Qy   5 LSILLLLSCSVLDSFGELIPQ 25
       |: |:||||: :| | ||
Db   8 LTALVLLSCNAICSLGCDLPQ 28

Query Match        31.3%;  Score 46;  DB 1;  Length 189;
Best Local Similarity  47.6%;  Pred. No. 12;
Matches   10;  Conservative  5;  Mismatches  6;  Indels  0;  Gaps  0;
```

CC chemical compound, or its metabolite, will be converted into a
CC carcinogenic or mutagenic form through metabolism in the liver.
SQ Sequence   490 AA;

Query Match   31.0%; Score 45.5; DB 1; Length 490;
Best Local Similarity   38.2%; Pred. No. 43;
Matches   13; Conservative   6; Mismatches   8; Indels   7; Gaps   1;

Qy 1 MDCQLSILILLSCSVLDSF------GELIPQPS 27
        | :|::|| ||| :|:| |:| ||:|
Db 1 MDSIVSLVLCLSCLILLSLWRQSSGRGKLPPGPT 34

RESULT 11
ID R81465 standard; Protein; 490 AA.
AC R81465;
DT 01-AUG-1996 (first entry)
DE Human derived cytochrome P450 2C9.
PT Human derived cytochrome; P450 2C9; commercial cDNA library; yeast;
PT transfection; recombinant production; expression vector; mammal;
KW immunisation; sensitisation; antibody; determination; detection;
KW non-cross reactive.
OS Homo sapiens.
PN J08027196-A.
PD 30-JAN-1996.
PF 13-JUL-1994; 161551.
PR 13-JUL-1994; JP-161551.
PA (SUMO) SUMITOMO CHEM CO LTD.
DR WPI; 96-136337/14.
DR N-PSDB; T17404.
PT Antibody recognising human derived cytochrome P450 2C9 - allows
PS specific detection of cytochrome P450 species in humans
PT Example 1; Pages 11-13; 13pp; Japanese.
CC The present sequence is the human derived cytochrome
CC P450 2C9, which was obtd. from a commercial cDNA library. Yeast
CC were transfected with an expression vector contg. the HDC cDNA,
CC cultured and then disrupted to give a microsomal fraction. The
CC HDC was purified from the fraction, and used to immunise and
CC sensitise a mammal. Blood was drawn from the mammal, and an
CC anti-HDC antibody isolated. The antibody obtd. recognises HDC
CC P450 2C9, partic. at a serum dilution rate of 1:10000, and is
CC substantially without cross reaction to other HDC P450 spp. .
SQ Sequence   490 AA;

Query Match   31.0%; Score 45.5; DB 1; Length 490;
Best Local Similarity   38.2%; Pred. No. 43;
Matches   13; Conservative   6; Mismatches   8; Indels   7; Gaps   1;

Qy 1 MDCQLSILILLSCSVLDSF------GELIPQPS 27
        | :|::|| ||| :|:| |:| ||:|
Db 1 MDSIVSLVLCLSCLILLSLWRQSSGRGKLPPGPT 34

RESULT 12
ID R93168 standard; Protein; 490 AA.
AC R93168;
DT 11-OCT-1996 (first entry)
DE Human cytochrome P450 molecular species 2C9 protein.
KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
KW evaluation; safety; fusion protein; metabolite; detoxification;
OS Homo sapiens.
KW carcinogenic.
PN J0805695-A.
PD 05-MAR-1996.
PF 15-JUL-1994; 164184.
PR 20-JUL-1993; JP-201120.
PR 30-JUL-1993; JP-208279.
PR 17-JUN-1994; JP-136053.

PA (SUMO) SUMITOMO CHEM CO LTD.
DR WPI; 96-182311/19.
DR N-PSDB; T28381.
PT Novel method for the evaluation of the safety of a cpd. - using a
PT human cytochrome P450 and yeast NADPH reductase to determine whether
PS the analyte cpd. is detoxified or metabolised to a carcinogen
PT Example 1; Page 20-22; 74pp; Japanese.
CC This is the amino acid sequence of the human cytochrome P450 molecular
CC species 2C9 protein. The corresp. gene was amplified from a human liver
CC derived cDNA library as 2 fragments of 0.9 and 0.6 kb using primers
CC to generate plasmid p229 for prodn. of the cytochrome only or into the
CC vector pAHRR to generate the plasmid p2C9R for co-prodn. with the yeast
CC NADPH-P450 reductase. The sequence is placed under control of the yeast
CC ADH gene promoter and terminator.
CC The vectors are used in a method for evaluating the safety of a cpd. by
CC reacting the test cpd. with recombinantly produced human cytochrome P450
CC mol. species 1A2 (T28380), 2C9, 2E1 (T28382), 3A4 (T28383) or their
CC variants (T28384-98) together with yeast NADPH-P450 reductase (either as
CC a fused protein or as a cell extract) and analysing the resultant
CC metabolite. The cpd. is considered "safe" if it is detoxified or not
CC rendered carcinogenic or "unsafe" if it is not detoxified or is
CC metabolised to a carcinogenic cpd.
SQ Sequence   490 AA;

Query Match   31.0%; Score 45.5; DB 1; Length 490;
Best Local Similarity   38.2%; Pred. No. 43;
Matches   13; Conservative   6; Mismatches   8; Indels   7; Gaps   1;

Qy 1 MDCQLSILILLSCSVLDSF------GELIPQPS 27
        | :|::|| ||| :|:| |:| ||:|
Db 1 MDSIVSLVLCLSCLILLSLWRQSSGRGKLPPGPT 34

RESULT 13
ID W27850 standard; Protein; 114 AA.
AC W27850;
DT 21-JUL-1998 (first entry)
DE Staphylococcus aureus protein of unknown function.
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
OS Staphylococcus aureus.
FH Key   Location/Qualifiers
FT Misc_difference   111
FT   /note= "not specified"
FT Misc_difference   112
FT   /note= "not specified"
PN WO9730070-A1.
PD 21-AUG-1997.
PF 19-FEB-1997; U02318.
PR 20-FEB-1996; US-011888.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
DR WPI; 97-424969/39.
DR N-PSDB; T83813.
PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against S.
PT aureus infection
PT Claim 6; Pages 309-310; 989pp; English.
CC The present sequence represents a Staphylococcus aureus protein of
CC unknown function. The DNA sequence was isolated from a library of
CC clones of S. aureus WCUH 29 in Escherichia coli. The DNA sequence can
CC be used in the construction of ribozymes and antisense sequences to
CC control the expression of Staphylococcus genes. The DNA sequence is
CC also useful as a source of regulatory elements for the control of
CC bacterial gene expression. The present protein may be used to produce
CC vaccines to enable a host to produce specific antibodies with
CC antibacterial action. These vaccines and antibodies would protect

```
CC   a host against invasion by S. aureus, and conditions relating to
CC   Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC   skin syndrome, and toxic shock syndrome.
SQ   Sequence   114 AA.

Qy   12  SCSVLDSFGELIPQ 25
         | |:|| || :|:|
Db   97  SASILDIFGRLLPR 110

Query Match      30.6%;  Score 45;  DB 1;  Length 114;
Best Local Similarity   57.1%;  Pred. No. 9.6;
Matches   8;  Conservative   3;  Mismatches   3;  Indels   0;  Gaps   0;


RESULT   14
R55214
ID   R55214  standard; Protein; 128 AA.
AC   R55214;
DT   01-FEB-1995  (first entry)
DE   Murine variable region light chain from 2B5.
KW   B cell lymphoma chimeric antibody; CD20; peripheral blood cells;
KW   cell lysis.
OS   Mus musculus.
PN   WO9411026-A.
PD   26-MAY-1994.
PF   12-NOV-1993;  U10953.
PR   13-NOV-1992;  US-978891.
PR   03-NOV-1993;  US-149099.
PA   (IDEC-) IDEC PHARM CORP.
PI   Anderson DR, Hanna N, Leonard JE, Newman RA, Rastetter WH;
PI   Reffme;
DR   N-PSDB;  Q65530.
DR   WPI;  94-183162/22.
PT   Treating B cell lymphoma with chimeric antibody - against CD20,
PT   causing rapid depletion of peripheral B cells, also new
PT   antibodies and hybridomas
PS   Disclosure; Fig 4; 10lpp; English.
CC   The sequence is the murine variable region light chain derived from
CC   murine anti-CD20 monoclonal antibody 2B5.
CC   See also Q65629-35.
SQ   Sequence   128 AA.

Qy    1  MDCQLSLI--LLLSCSVLDSFGELI 23
         |:|::| |::|| | |:::
Db    1  MDFQVQIISFLLISASVIMSRQIV 25

Query Match      30.6%;  Score 45;  DB 1;  Length 128;
Best Local Similarity   44.0%;  Pred. No. 11;
Matches   11;  Conservative   7;  Mismatches   5;  Indels   2;  Gaps   1;


RESULT   15
W41568
ID   W41568  standard; Protein; 908 AA.
AC   W41568;
DT   22-JUN-1998  (first entry)
DE   Human metabotropic glutamate receptor.
KW   Metabotropic glutamate receptor; mGluR8; human; agonist; antagonist;
KW   neurodegenerative disease; antipsychotic; anticonvulsant; analgesic;
KW   anxiolytic; antidepressant; antiemetic; therapy.
OS   Homo sapiens.
PN   EP-816498-A2.
PD   07-JAN-1998.
PF   02-JUL-1997;  304821.
PR   03-JUL-1996;  US-021243.
PA   (ELIL ) LILLY & CO ELI.
PI   Belagaje RM, Wu S;
DR   WPI;  98-054913/06.
DR   N-PSDB;  V04206-07.
PT   Human metabotropic glutamate receptor protein - agonists and
PT   antagonists of which are useful to treat neurodegenerative diseases
PS   Claim 1; Page 46-48; 56pp; English.
CC   This protein comprises a novel human metabotropic glutamate
CC   receptor, designated mGluR8.  Its amino acid sequence was deduced
CC   from cDNA clones isolated from a human foetal retina cDNA library.
CC   The metabotropic glutamate receptors are linked to mutiple second
CC   messenger pathways.  They function to modulate the presynaptic
CC   release of glutamate, and the postsynaptic sensitivity of the
CC   neuronal cell to glutamate excitation.  Host cells transfected with
CC   an expression vector comprising nucleic acids encoding mGluR8,
CC   especially RGT-18 host cells transfected with pGT-h-mGluR8, are
CC   claimed.  They can be used in claimed methods for evaluating the
CC   effectiveness of a test compound for the treatment or prevention of
CC   a condition associated with a deficiency or excess of stimulation
CC   of the human mGluR8 receptor. mGluR8 agonists and antagonists can
CC   be used to treat neurodegenerative diseases, or as antipsychotic,
CC   anticonvulsant, analgesic, anxiolytic, antidepressant and
CC   antiemetic agents.
SQ   Sequence   908 AA.

Qy    3  CQL--SILLLSCSVLDSFGELIPQPSNE 29
         | |  |||::::||| |  :|: ||
Db   751  CSLGYSILLMVTCTVYANKTRGVPETFNE 779

Query Match      30.6%;  Score 45;  DB 1;  Length 908;
Best Local Similarity   37.9%;  Pred. No. 1e+02;
Matches   11;  Conservative   7;  Mismatches   9;  Indels   2;  Gaps   1;

Search completed: May   9, 2000, 21:52:05
Job time: 3757 sec
```

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            May 9, 2000, 22:31:13 ; Search time 71.69 Seconds
                                           (without alignments)
                                           262.094 Million cell updates/sec

Title:             US-09-104-340-4
Perfect score:     1466
Sequence:          1 MDCQLSILILLSCSVLDSFG..........VPIGKCSCNAGYEERGPMCQ 271

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters:  225878
Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :       SPTREMBL_12:*
              1: sp_archea:*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	977	66.6	985	13	Q91694	Q91694 xenopus lae
2	977	66.6	986	13	Q91845	Q91845 xenopus lae
3	934.5	63.7	993	13	042422	042422 gallus gall
4	851.5	58.1	981	13	013146	013146 brachydanio
5	838.5	57.2	981	13	073879	073879 brachydanio
6	768.5	52.4	985	13	Q91571	Q91571 xenopus lae
7	767	52.3	1055	4	043477	043477 homo sapien
8	739.5	50.4	984	4	043569	043569 homo sapien
9	739.5	50.4	984	4	095142	095142 homo sapien
10	724.5	49.4	974	13	095143	095143 xenopus lae
11	713.5	48.7	973	4	091735	091735 xenopus lae
12	679	46.3	988	13	007498	007498 homo sapien
13	594.5	40.6	952	13	007494	007494 gallus gall
14	589	40.2	938	13	060669	060669 mus musculu
15	563.5	38.4	902	13	091736	091736 xenopus lae
16	534	36.4	976	13	073878	073878 brachydanio
17	528.5	36.1	1006	4	015197	015197 homo sapien
18	524.5	35.8	1014	11	008644	008644 mus musculu
19	407.5	27.8	1035	5	Q9xzl6	Q9xzl6 drosophila
20	407.5	27.8	1080	5	Q9yij0	Q9yij0 drosophila
21	397.5	27.1	1019	5	096435	096435 drosophila
22	333.5	22.7	871	13	073875	073875 brachydanio
23	280	19.1	815	13	057458	057458 xenopus lae
24	250.5	17.1	1122	5	061460	061460 caenorhabdi
25	220	15.0	893	5	Q9y1y3	Q9y1y3 ephydatia f
26	194.5	13.3	73	13	Q91626	Q91626 xenopus lae
27	97	6.6	413	2	005989	005989 staphylococ
28	96	6.5	698	10	023268	023268 arabidopsis
29	90	6.1	3110	4	093022	093022 homo sapien
30	89.5	6.1	668	10	055616	055616 arabidopsis
31	87.5	6.0	788	13	007012	007012 xenopus lae
32	86.5	5.9	305	2	032378	032378 methylobaci
33	86	5.9	919	5	021477	021477 caenorhabdi
34	85.5	5.8	347	10	018395	018395 caenorhabdi
35	84.5	5.8	714	5	015718	015718 dictyosteli
36	84.5	5.8	865	10	023617	023617 arabidopsis
37	84	5.7	701	10	096337	096337 brassica na
38	84	5.7	784	4	015495	015495 homo sapien
39	83	5.7	784	4	096206	096206 plasmodium
40	82.5	5.6	1817	5	050966	050966 borrelia bu
41	82	5.6	369	2	097702	097702 canis famil
42	82	5.6	784	6	016366	016366 caenorhabdi
43	81.5	5.6	630	5	022728	022728 arabidopsis
44	81	5.5	889	10	025431	025431 helicobacte
45	81	5.5	573	2	Q22747	Q22747 caenorhabdi

ALIGNMENTS

RESULT 1
Q91694

```
ID   Q91694        PRELIMINARY;   PRT;   985 AA.
AC   Q91694;
DT   01-NOV-1996 (TREMBLrel. 01, Created)
DT   01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR PAG PRECURSOR (EC 2.7.1.112)
DE   (PAGLIACCIO).
GN   PAG.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NEURAL CREST;
RX   MEDLINE; 95001564.
RA   WINNING R.S., SARGENT T.D.;
RT   "Pagliaccio, a member of the Eph family of receptor tyrosine kinase
RT   genes, has localized expression in a subset of neural crest and neural
RT   tissues in Xenopus laevis embryos.";
RL   Mech. Dev. 46:219-229(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY
CC       SIMILARITY). MAY PLAY A ROLE IN THE DIFFERENTIATION OF CRANIAL
CC       NEURAL CREST AND OTHER TISSUES.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: LOCALIZED EXPRESSION IN A SUBSET OF NEURAL
CC       CREST AND NEURAL TISSUES IN EMBRYOS.
CC   -!- DEVELOPMENTAL STAGE: PRESENT TRANSIENTLY IN VISCERAL ARCH 3. ALSO
CC       EXPRESSED IN THE FOREBRAIN, RHOMBOMERES R3 AND R5 OF THE HINDBRAIN
CC       AND IN THE PRONEPHROS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; L26099; AAA64464.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V; 1.
```

```
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    985       TYROSINE-PROTEIN KINASE RECEPTOR PAG.
FT   DOMAIN       21    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    985       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       66    115       IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN      191    325       CYS-RICH.
FT   DOMAIN      326    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    533       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      620    881       PROTEIN KINASE.
FT   NP_BIND     626    634       ATP (BY SIMILARITY).
FT   DISULFID     73    108       BY SIMILARITY.
FT   BINDING     652    652       ATP (BY SIMILARITY).
FT   ACT_SITE    745    745       BY SIMILARITY.
FT   MOD_RES     778    778       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    407    407       POTENTIAL.
SQ   SEQUENCE    985 AA;  109729 MW;  026BA8A5 CRC32;

Query Match             66.6%;  Score 977;  DB 13;  Length 985;
Best Local Similarity   70.6%;  Pred. No. 2.4e-83;
Matches  175;  Conservative  39;  Mismatches  32;  Indels  2;  Gaps  2;

Qy    26 PSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
         |:||| |||||||||||||| | | |||||||::|| |||||||||||||||: ||||||||
Db    27 PASEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMESSQNNWLRT 86

Qy    85 NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETENLYYMESDDDHGVKFREHQFTKID 144
         :|:||:|:|| |||||||||||||||:||:|||||||||||:|||| |:| ||||:|||||
Db    87 DWIPRSGAQRYVEIKFTLRDCNSLPGVMGTCKETENLYYESNNDKERFIRETQYVKID 146

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPF 204
         ||||||||||:|:|||||:||||||:|||:|:|||||||||||||||:||||||:|:|||:
Db   147 TIAADESFTQVDIGDRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPL 206

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDVPKMYCGADGEWLVPIGNCLCNAGY 263
         ||:|||:|||:|| |:|:||||||||||:||:|| :|||||:||||||||||||:|||||:
Db   207 TVRNLAQFPDITGSDTSSLVEVRGSCVDNSEEKDVPKMYCGADGEWLVPIGNCLCNAGF 266

Qy   264 EERGFMCQ 271
         |||:| ||
Db   267 EEHNGGCQ 274

RESULT  2
Q91845

ID   Q91845           PRELIMINARY;    PRT;   986 AA.
AC   Q91845;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TYROSINE-PROTEIN KINASE RECEPTOR SEK-1 PRECURSOR (EC 2.7.1.112).
GN   SEK-1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96125143.
RA   XU Q., ALLDUS G., HOLDER N., WILKINSON D.G.;
RT   "Expression of truncated Sek-1 receptor tyrosine kinase disrupts the
RT   segmental restriction of gene expression in the Xenopus and zebrafish
RT   hindbrain.";
RL   Development 121:4005-4016(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY (BY
CC       SIMILARITY). REQUIRED FOR INTERACTIONS THAT REGULATE THE SEGMENTAL
CC       IDENTITY OR MOVEMENT OF CELLS.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSION OCCURS IN R3, R5 AND TRANSIENTLY
CC       AT LOWER LEVELS IN R2.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPH FAMILY.
DR   EMBL; X91191; CAA62601.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTPEIII.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Immunoglobulin domain.
FT   SIGNAL        1     15       POTENTIAL.
FT   CHAIN        16    986       TYROSINE-PROTEIN KINASE RECEPTOR SEK-1.
FT   DOMAIN       16    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       66    115       IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN      191    325       CYS-RICH.
FT   DOMAIN      326    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    533       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      620    881       PROTEIN KINASE.
FT   NP_BIND     626    634       ATP (BY SIMILARITY).
FT   DISULFID     73    108       BY SIMILARITY.
FT   BINDING     652    652       ATP (BY SIMILARITY).
FT   ACT_SITE    745    745       BY SIMILARITY.
FT   MOD_RES     778    778       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    407    407       POTENTIAL.
SQ   SEQUENCE    986 AA;  109840 MW;  028DBBDA CRC32;

Query Match             66.6%;  Score 977;  DB 13;  Length 986;
Best Local Similarity   70.6%;  Pred. No. 2.4e-83;
Matches  175;  Conservative  39;  Mismatches  32;  Indels  2;  Gaps  2;

Qy    26 PSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
         |:||| |||||||||||||| | | |||||||::|| |||||||||||||||: ||||||||
Db    27 PASEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMESSQNNWLRT 86

Qy    85 NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETENLYYMESDDDHGVKFREHQFTKID 144
         :|:||:|:|| |||||||||||||||:||:|||||||||||:|||| |:| ||||:|||||
Db    87 DWIPRSGAQRYVEIKFTLRDCNSLPGVMGTCKETENLYYESNNDKERFIRETQYVKID 146

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPF 204
         ||||||||||:|:|||||:||||||:|||:|:|||||||||||||||:||||||:|:|||:
Db   147 TIAADESFTQVDIGDRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPL 206

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDVPKMYCGADGEWLVPIGNCLCNAGY 263
         ||:|||:|||:|| |:|:||||||||||:||:|| :|||||:||||||||||||:|||||:
Db   207 TVRNLAQFPDITGSDTSSLVEVRGSCVDNSEEKDVPKMYCGADGEWLVPIGNCLCNAGF 266

Qy   264 EERGFMCQ 271
         |||:| ||
Db   267 EEHNGGCQ 274
```

RESULT 3

ID   O42422   PRELIMINARY;   PRT;   993 AA.
AC   O42422;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE RECEPTOR CEPHA7).
DE   PROTEIN KINASE RECEPTOR TYROSINE KINASE RECEPTOR CEPHA7 PRECURSOR (EC 2.7.1.112) (TYROSINE-
GN   CEPHA7.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 98092111.
RA   ARAUJO M., NIETO M.A.;
RT   "The expression of chick EphA7 during segmentation of the central and
RT   peripheral nervous system.";
RL   Mech. Dev. 68:173-177(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WITHIN THE NERVOUS SYSTEM, EXPRESSION IS
CC   RESTRICTED TO PROSOMERES 1 AND 2 IN THE DIENCEPHALON AND ALL THE
CC   RHOMBOMERES IN THE HINDBRAIN DURING SEGMENTATION STAGES. LATER ON,
CC   A SUPERIMPOSED PATTERN APPEARS THAT CORRELATES WITH THE FORMATION
CC   OF SEVERAL AXONAL TRACTS. IN THE SOMITIC MESODERM, THE EXPRESSION
CC   CORRELATES WITH SEGMENTATION AND THE GUIDANCE OF BOTH NEURAL CREST
CC   AND MOTOR AXONS THROUGH THE SCLEROTOMES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; Y14271; CAA74643.1; -.
DR   HSSP; P00523; 2PTK.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PRINTS; PR00014; FNTPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   30   POTENTIAL.
FT   CHAIN   31   993   EPH-LIKE RECEPTOR TYROSINE KINASE.
FT   DOMAIN   31   551   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   552   572   POTENTIAL.
FT   DOMAIN   573   993   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   67   116   IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN   192   328   CYS-RICH.
FT   DOMAIN   329   438   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   439   536   FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN   628   889   PROTEIN KINASE.
FT   DOMAIN   634   642   ATP (BY SIMILARITY).
FT   NP_BIND   74   109   ATP (BY SIMILARITY).
FT   BINDING   660   660   BY SIMILARITY.
FT   DISULFID   753   786   BY SIMILARITY.
FT   ACT_SITE   753   753   BY SIMILARITY.
FT   BINDING   660   660   ATP (BY SIMILARITY).
FT   MOD_RES   786   786   PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD   343   343   POTENTIAL.
FT   CARBOHYD   410   410   POTENTIAL.
SQ   SEQUENCE   993 AA; 111366 MW; 6B6BFD17 CRC32;

Query Match   63.7%; Score 934.5; DB 13; Length 993;

Qy   10 LLSCSV----LDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHYTPI 65
         :: ||| | | | | || ||||| | |||| || | |||||||| ::| ||||
Db   11 IMLCSWLLRFAHTGE--AQAAKEVILLDSKAQOTELEWISSPPNGWEEISGLDENYTPI 68

Qy   66 RIYQVCNVMDHSQNNWLRTNWVPRNSAQKIIYELKFTLRDCNSIPLVLGTCKETFNLYYM 125
         |||||| ||:|| || ::|||||||||| |:|||||||||||||||||:|| ||||||||||
Db   69 RIYQVCQVMESNQNNWLRTNWIAKSNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLYYY 128

Qy 126 ESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQD 185
         |||| ::| ||: |: | ||| || | || | |::|: ||||:|| |||:|:|||:||
Db 129 ETDYDTGRNIRENQYVKIDTIAADESFTQGGDLGERKMLNTEVREIGPLSKKGFYLAFQD 188

Qy 186 VGACVALVSVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEYRGSCVNNSKE--EDPPRM 242
         || |:| |||| :||||||| | ||| ||| | | |||||| | || :| |:| |
Db 189 VGACIALVSVKVTYKKCWSIIENLAIFPDTVTGSEFSSLVEVRAGTICVSSAAEEEANSPKM 248

Qy 243 YCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
         :|| |||||||| |||||||||||| | |:
Db 249 HCSAEGEWLVPIGKCICKAGYQQKGDTCE 277

RESULT 4

Best Local Similarity 65.4%; Pred. No. 2.4e-79;
Matches 176; Conservative 43; Mismatches 41; Indels 9; Gaps 4;

ID   O13146   PRELIMINARY;   PRT;   981 AA.
AC   O13146;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE KINASE 1 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN KINASE
DE   RECEPTOR ZEK1).
GN   ZEK1.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC   Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC   Cyprinoidea; Cyprinidae; Rasborinae; Danio.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97329578.
RA   BOVENKAMP D.E., GREER P.;
RT   "Novel Eph-family receptor tyrosine kinase is widely expressed in the
RT   developing zebrafish nervous system.";
RL   Dyn. 209:166-181(1997).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC   ROLE IN EARLY PATTERN FORMATION WITHIN THE DEVELOPING NERVOUS
CC   SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN THE DEVELOPING ZEBRAFISH
CC   NERVOUS SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U89295; AAC60220.1; -.
DR   HSSP; P00523; 2PTK.
DR   ZFIN; ZDB-GENE-990415-58; zek1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_2; FALSE_NEG.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL   1   20   BY SIMILARITY.

Left column

```
FT   CHAIN        21    981       EPH-LIKE KINASE 1.
FT   DOMAIN       21    981       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    546    566       POTENTIAL.
FT   DOMAIN      567    981       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       64    113       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      192    325       CYS-RICH.
FT   DOMAIN      326    438       FIBRONECTIN TYPE-III.
FT   DOMAIN      439    532       FIBRONECTIN TYPE-III.
FT   DOMAIN      626    887       PROTEIN KINASE.
FT   NP_BIND     632    640       ATP (BY SIMILARITY).
FT   DISULFID     71    106       BY SIMILARITY.
FT   BINDING     658    658       ATP (BY SIMILARITY).
FT   ACT_SITE    751    751       BY SIMILARITY.
FT   MOD_RES     784    784       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
FT   CARBOHYD    435    435       POTENTIAL.
FT   CARBOHYD    485    485       POTENTIAL.
FT   VARIANT     141    141       S -> P.
FT   VARIANT     141    141       S -> N.
SQ   SEQUENCE   981 AA;  109654 MW;  10D38182 CRC32;
```

```
Query Match             58.1%;  Score 851.5;  DB 13;  Length 981;
Best Local Similarity   58.6%;  Pred. No. 1.4e-71;
Matches  156;  Conservative   44;  Mismatches   57;  Indels    9;  Gaps    2;

Qy    9 LLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHWEEISGVDEHYTPIRTY 68
        ||:| :|| ||| :|                 ::||:|::|  ||||| |:|:| |||
Db   15 ILVLCQALRNY------PDNEVLLLDSMSAPGDLGWEAPSEGWEEISVMDERNIPMRIY 68

Qy   69 QVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPLVLGTCKEINLYYMESD 128
        ||||||:: ||||||| ::: | :| ::|:||||||||:| ::||||| |:|||:||:
Db   69 QVCNVMEANQNNMLRTGLIQREGAQRVVEIKFTLRDCNSLPGVPGTCKETFNVYYHESN 128

Qy  129 DDHGV---KFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQD 185
        :|||    :| |. |:|||||||||||| :|:|||::|||||:|::   ||.|.|||||
Db  129 NAVAAPLRHIRESQYIKIDTIAADESFQTDVGDRVMKLNTEVRDISGLSKRGLYLAFQD 188

Qy  186 VGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPPRMYCS 245
        :|||:||:|:||:|:.||:: ::||: ||||  | :||||||:||:: :||| ||| ||
Db  189 LGACIALVSVRVYTKRCPLAVLNLARFPDTVTGGDSALVEVRGTCVEDAEELEGPRMFCS 248

Qy  246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
        :||||||:|:||.:|| :||| :|
Db  249 ADGGWLVPIGRCVCRPGFEEVDGHCQ 274
```

```
RESULT 5
073879
ID   073879      PRELIMINARY;  PRT;  880 AA.
AC   073879;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE RTK4 (FRAGMENT).
GN   RTK4.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC   Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC   Cyprinoidea; Cyprinidae; Rasborinae; Danio.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   COOKE J.E.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   COOKE J.E., XU Q., WILSON S.W., HOLDER N.;
RL   Dev. Genes Evol. 206:515-531(1997).
DR   EMBL; AJ005030; CAA06303.1; -.
DR   HSSP; P00523; 2PTK.
DR   ZFIN; ZDB-GENE-990415-61; rtk4.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
```

Right column

```
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00041; fn3; 2.
FT   NON_TER       1      1
FT   NON_TER     880    880
SQ   SEQUENCE   880 AA;  98862 MW;  7C139606 CRC32;
```

```
Query Match             57.2%;  Score 838.5;  DB 13;  Length 880;
Best Local Similarity   68.0%;  Pred. No. 2.1e-70;
Matches  149;  Conservative   36;  Mismatches   33;  Indels    1;  Gaps    1;

Qy   54 EISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPLVL 113
        |:| :|| |||||:||||||||::||||||||.|:: | .|:|:|::|:||||||:| :|
Db    1 EVSIMDEKNIPIRIYQVCNVMEPSQNNWLRTHWIQRGAAQRIYIEIKFTLRDCNSLPGVI 60

Qy  114 GTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGP 173
        ||||||||||| ||.|::  ::.| .||:|||.|||||||| |:|||.:||||||:|.||
Db   61 GTCKETFNLYLESDSDNERYAHESRFSKIDTVAADESFTQVDIGDRIMKLNTEVRDGV 120

Qy  233 NSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
        .|:||:.. ||| .:|||||||| |.||.|||||.. ||
Db  181 HSEEQEVPKMYCGADGEWLVPIGNCLCNPGYEERNEQCQ 219
```

```
RESULT 6
Q91571
ID   Q91571      PRELIMINARY;  PRT;  985 AA.
AC   Q91571;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR XEK PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR XEK).
GN   XEK.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95215070.
RA   JONES T.L., KARAVANOVA I., MAENO M., ONG R.C., KUNG H.-F., DAAR I.O.;
RT   "Expression of an amphibian homolog of the Eph family of receptor
RT   tyrosine kinases is developmentally regulated.";
RL   Oncogene 10:1111-1117(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
CC       ROLE IN THE DEVELOPMENT OR FUNCTION OF THE CENTRAL NERVOUS SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED, IT DECREASES AT MID
CC       BLASTULA TRANSITION AND REAPPEARS AT LATE NEURULATION. EXPRESSED
CC       AT HIGHER LEVELS IN THE ANTERIOR AND DORSAL REGIONS OF EMBRYONIC
CC       STAGES 16, 24 AND 37. IN ADULT IT APPEARS TO BE UBIQUITOUSLY
CC       EXPRESSED WITH HIGHER EXPRESSION IN BRAIN AND OVARY. EXPRESSION IN
CC       THE BRAIN, BRACHIAL ARCHES, TRIGEMINAL FACIAL GANGLION, AND THE
CC       RETINA OF SWIMMING TADPOLE STAGE OF DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; U14164; AAA74888.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
```

Left column

```
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; FALSE_NEG.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    985       EPHRIN TYPE-B RECEPTOR XEK.
FT   DOMAIN       6     12       POLY-LEU.
FT   DOMAIN      20    542       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   543    563       POTENTIAL.
FT   DOMAIN     564    985       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      56    105       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN     185    321       CYS-RICH.
FT   DOMAIN     322    431       FIBRONECTIN TYPE-III.
FT   DOMAIN     432    529       FIBRONECTIN TYPE-III.
FT   DOMAIN     620    883       PROTEIN KINASE.
FT   NP_BIND    626    634       ATP (BY SIMILARITY).
FT   DISULFID    63     98       BY SIMILARITY.
FT   BINDING    652    652       ATP (BY SIMILARITY).
FT   ACT_SITE   745    745       BY SIMILARITY.
FT   MOD_RES    779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD   336    336       POTENTIAL.
FT   CARBOHYD   428    428       POTENTIAL.
FT   CARBOHYD   482    482       POTENTIAL.
SQ   SEQUENCE   985 AA;  110104 MW;  2416A8F9 CRC32;
```

```
Query Match             52.4%;
Best Local Similarity   52.0%;   Pred. No. 8.9e-64;   DB 13;   Length 985;
Matches  143;   Conservative  55;   Mismatches  64;   Indels  13;   Gaps  5;

Qy     4  QLSILILLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGGWEEISGVDEHYT  63
          |||||||
Db     2  ELNVLLILLCLSGGQVGAV-----EETLMDTRTATAELGWTANPSSGGWEEVSGYDENLN  55

Qy    64  PIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLX  123
          |||||||||
Db    56  TIRTYQVCNVFPGKQNNWLLTFIPIRRGAHRVVEMRFTVRDCSSLPNVPGSCKETFNLX  115

Qy   124  YMESDDDHGVK-----FREHQFTKIDTIAADESETQMDLGDRILKLNTEIREVGPVNKKGF  179
          |:|:| :|
Db   116  YYETDSNIENKISTFWNESPYLKVDTIAADESFSQVDEGGRLMKVNTEVRSFGPLTRSGF  175

Qy   180  YLAFQDVGACVALVSVRVYFFKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEED  238
Db   176  YLAFQDVGACMSLLLSVRVFFKEMPSVVQNLLVFPETMVGAESTSLVIARGTCIPNAEEVD  235

Qy   239  PP-RMYCSTEGEWLVPIGKCSCNAGYE--ERGFMCQ  271
Db   236  VPIKLYCNGDGEMMVPIGKCTCKAGYEPENHVVCK  270
```

```
RESULT 7
ID   O43477          PRELIMINARY;   PRT;   1055 AA.
AC   O43477;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTEIN-TYROSINE KINASE EPHB2.
GN   EPHB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96154673.
RA   IKEGAKI N., TANG X.X., LIU X.G., BIEGEL J.A., ALLEN C., YOSHIOKA A.,
RA   SULMAN E.P., BRODEUR G.M., PLEASURE D.E.;
```

Right column

```
RT   "Molecular characterization and chromosomal localization of DRT
RT   (EPHT3): a developmentally regulated human protein-tyrosine kinase
RT   gene of the EPH family.";
RL   Hum. Mol. Genet. 4:2033-2045(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   TANG X.X., PLEASURE D.E., IKEGAKI N.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF025304; AAB94602.1; -.
RP   SEQUENCE FROM N.A.
DR   HSSP; P05523; 2PTK.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; EPHTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Tyrosine-protein kinase.
SQ   SEQUENCE   1055 AA;  117492 MW;  EAF28155 CRC32;
```

```
Query Match             52.3%;
Best Local Similarity   55.0%;   Pred. No. 1.3e-63;   DB 4;   Length 1055;
Matches  138;   Conservative  52;   Mismatches  53;   Indels  8;   Gaps  4;

Qy    29  EVNLLDSKTIQGELGWISYPSHGGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVP  88
          | :||         |||||||  :|||          ||||||||||          ::||
Db    20  EETLMDSTTATAELGWMYHPPSGWEEVSGYDENMNTIRTYQVCNVNFESSQNNWLRTKFIR  79

Qy    89  RNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK-----FREHQFTKID  144
          |::|  |: ||| :||||:|||:: | |||||||:||: :     :|     |||||||| |
Db    80  RRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLIYYEADFDSATKIFPNWMENPWVKVD  139

Qy   145  TIAADESFTQMDLGDRILKLNTEIREVGPVNKGFYLAFQDVGACVALVSVRVYFKKCPF  204
          ||||||| :|:|||  |:|:|||:|  :|||::|:|||||  |||  |:::||| ||:|||
Db   140  TIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAPQDYGGCMSLIAVRVFKYRKCPR  199

Qy   205  TVKNIAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAG  262
          :: | | |   :| |:::||| ||||| :: |:|| || ::|| |: ||||||:| | ||
Db   200  IIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGDGEWMVPIGRCMCKAG  259

Qy   263  YE--ERGFMCQ  271
          :|  :| | :|
Db   260  FEAVENGTVCR  270
```

```
RESULT 8
ID   O43569          PRELIMINARY;   PRT;   943 AA.
AC   O43569;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1C
DE   (EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1D).
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   STEIN E., SCHOECKLMANN H.O., DANIEL T.O.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF037333; AAB94627.1; -.
DR   EMBL; AF037333; AAB94627.1; -.
DR   EMBL; AF037334; AAB94628.1; -.
DR   HSSP; P08631; 2HCK.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 2.
DR   PFAM; PF00536; SAM; 1.
```

```
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
     PRINTS; PR00014; FNTYPEIII.
SQ   SEQUENCE  943 AA; 105291 MW; C4C78A21 CRC32;

Qy   Query Match    50.4%;  Score 739.5;  DB 4;  Length 943;
     Best Local Similarity  50.7%;  Pred. No. 4.4e-61;
     Matches  138;  Conservative  53;  Mismatches  64;  Indels  17;  Gaps  5;

Qy    7  ILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIR  66
         :||||| ||                  :|
Db    7  LILLLESAV------------AAMEETLMDTRIATAELGWTANPASGWEEVSGYDENLNTIR  56

Qy   67  TYQVCNVADHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME  126
Db   57  TYQVCNVFPNQNNWLLTFINRGAHRIYSEMRFTVRDCSSLPNVPGSCKETFNLYYYE  116

Qy  127  SDDDHGVK---FREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLA  182
Db  117  TDSVIATKKSAFWSEAPYLKVDTIAADESFSQVDFERGRLMKVNTEVRSFGPLIRNGFYLA  176

Qy  183  FQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-  241
Db  177  FQDYGACMILLSVRVFFKKCPSIVQNFAVPEPETMTGAESTSLVIARGTCIPNAEEVDVPI  236

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYE-ERGEMCQ  271
Db  237  KLYCNGDGEWMVPIGRCTCKPGYEPENSVACK  268
```

```
RESULT    9

ID   O95142           PRELIMINARY;       PRT;   984 AA.
AC   O95142;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH-LIKE RECEPTOR TYROSINE KINASE HEPHB1.
GN   EPHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   STEIN E., HUYNH-DO U., LANE A., CERRETTI D.P., DANIEL T.O.;
RT   "Nck recruitment to Eph receptor, EphB1/ELK, couples ligand activation
RT   to c-Jun kinase.";
RL   J. Biol. Chem. 0:0-0(1997).
DR   EMBL; AF037331; AAD02030.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW   Receptor; Kinase.
SQ   SEQUENCE  984 AA; 110041 MW; DF5C8ED3 CRC32;

Qy   Query Match    50.4%;  Score 739.5;  DB 4;  Length 984;
     Best Local Similarity  50.7%;  Pred. No. 4.7e-61;
     Matches  138;  Conservative  53;  Mismatches  64;  Indels  17;  Gaps  5;

Qy    7  ILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIR  66
Db    7  LILLLESAV------------AAMEETLMDTRTATAELGWTANPASGWEEVSGYDENLNTIR  56

Qy   67  TYQVCNVADHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYME  126
Db   57  TYQVCNVFPNQNNWLLTFINRGAHRIYSEMRFTVRDCSSLPNVPGSCKETFNLYYYE  116

Qy  127  SDDDHGVK---FREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLA  182
Db  117  TDSVIATKKSAFWSEAPYLKVDTIAADESFSQVDFERGRLMKVNTEVRSFGPLIRNGFYLA  176

Qy  183  FQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-  241
Db  177  FQDYGACMILLSVRVFFKKCPSIVQNFAVPEPETMTGAESTSLVIARGTCIPNAEEVDVPI  236

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYE-ERGEMCQ  271
Db  237  KLYCNGDGEWMVPIGRCTCKPGYEPENSVACK  268
```

```
RESULT   10

ID   Q91735           PRELIMINARY;       PRT;   974 AA.
AC   Q91735;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPHRIN TYPE-B RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR) (TCK).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96068901.
RA   SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;
RT   "Novel members of the eph receptor tyrosine kinase subfamily expressed
RT   during Xenopus development.";
RL   Oncogene 11:1745-1752(1995).
RN   [1]
RP   FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY (BY
     SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO IN PRE-SOMITIC
CC       MESODERM, CAUDAL SOMITES, MIDBRAIN, AND CEMENT GLAND. MOST
CC       ABUNDANT IN ADULT BRAIN, EYE, HEART, LUNG AND OVARY. LOWER LEVELS
CC       IN INTESTINE, KIDNEY, OVIDUCT AND PHARYNX.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE PROTEIN-TYROSINE RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L43620; AAA93526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17    974       EPHRIN TYPE-B RECEPTOR.
FT   DOMAIN      17    534       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   535    555       POTENTIAL.
FT   DOMAIN     556    974       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      53    102       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN     178    315       CYS-RICH.
FT   DOMAIN     316    423       FIBRONECTIN TYPE-III.
FT   DOMAIN     424    520       FIBRONECTIN TYPE-III.
FT   DOMAIN     609    872       PROTEIN KINASE.
FT   NP_BIND    615    623       ATP (BY SIMILARITY).
FT   DISULFID    60     95       BY SIMILARITY.
FT   BINDING    641    641       ATP (BY SIMILARITY).
FT   ACT_SITE   734    734       BY SIMILARITY.
FT   MOD_RES    768    768       PHOSPHORYLATION (AUTO-) (POTENTIAL).

Qy  183  FQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-  241
Db  177  FQDYGACMILLSVRVFFKKCPSIVQNFAVPEPETMTGAESTSLVIARGTCIPNAEEVDVPI  236

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYE-ERGEMCQ  271
Db  237  KLYCNGDGEWMVPIGRCTCKPGYEPENSVACK  268

Qy    7  ILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIR  66
Db    7  LILLLESAV------------AAMEETLMDTRTATAELGWTANPASGWEEVSGYDENLNTIR  56

Qy   67  TYQVCNVADHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYME  126
Db   57  TYQVCNVFPNQNNWLLTFINRGAHRIYSEMRFTVRDCSSLPNVPGSCKETFNLYYYE  116

Qy  127  SDDDHGVK---FREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLA  182
Db  117  TDSVIATKKSAFWSEAPYLKVDTIAADESFSQVDFERGRLMKVNTEVRSFGPLIRNGFYLA  176
```

```
FT   CARBOHYD       330    330    POTENTIAL.
FT   CARBOHYD       420    420    POTENTIAL.
SQ   SEQUENCE     974 AA;  108263 MW;  BD419F6D CRC32;

Query Match     49.4%;  Score 724.5;  DB 13;  Length 974;
Best Local Similarity   48.9%;  Pred. No. 1.2e-59;
Matches  134;  Conservative  51;  Mismatches  68;  Indels  21;  Gaps  6;

Qy     6 SILLLLSCSVLLDSFGELIPQPSNEVNLLDSKYIQGELGWISYPSHGWEEISGVDEHYTPI 65
         || ::|:| :|||||||||||| |::|:| |||||||||||||||||||||| ||
Db     4 ------------EEILMDTKWTISELAWVAYPDSGWEEVSGYDEASNPI 54

Qy    66 RTYQVCNVMDHSQONNWLRTNWVPRNSAQKIYELKFTLRDCNSIPLVLGTCKETFNLYYM 125
         ||||||||||||||||||||||| :|||||||||| ||||||||||||||||||| ::
Db    55 RTYQVCNVRDSNQONNWLRTQFIPRODVQRVVELKFTIVRDCNSLPNLRGSCKETFNFYY 114

Qy   126 ESDDDHGVK----FREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYL 181
         ::|:| | :|||||||| |:||| ||||| ::|| :||| ||
Db   115 ESDSDSASADSPFWMENPYIKVDTIAPDESFSRRDSG---RVNTKIRSFGPISRAGFYL 170

Qy   182 AFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP 240
         ||||||||| |:||||||:|:|||||||| :|||| |||||||||||||||||||||||
Db   171 AFQDLGACVSLISVRVFFKKCPRTAGFASFPETITGAEPISLVIAPGTCVPNALEVSVP 230

Qy   241 -RMYCSTEGEWLVPIGKCSCNAGYEERG--FMCQ 271
         :||||||| ||||||||||||||||| :|:|
Db   231 LKLYCNGDGDWMVPVGACTCAAGFEPAGKDTCQ 264
```

```
Query Match     48.8%;  Score 713.5;  DB 4;  Length 973;
Best Local Similarity   54.2%;  Pred. No. 1.3e-58;
Matches  129;  Conservative  47;  Mismatches  55;  Indels  7;  Gaps  4;

Qy    41 ELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQONNWLRTNWVPRNSAQKIYVELK 100
         ::||| :|:|:|||||||| |:|| |:|| :|:
Db    20 ELGWTANPASGWEEVSGYDENLNTIRIYQVCNVFEPNQONNWLLTTFINRRGAHRIYSEMR 79

Qy   101 FTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK----FREHQFTKIDTIAADESFTQMD 156
         |||||||||||| ||||||||||||||| :| |||||||| |:|||||||||:|
Db    80 FTVRDCSSLPNVPGSCKETFNLYYETDSVIATKKSAFWSEAPYLKVDTIAADESFSQVD 139

Qy   157 LGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTV 216
         :|| :||||| ::| |||| ||||||||||||||| |:|||||:|:|||||||||:
Db   140 FRGRLMVNTEVRSFGPLTRNGFYLAFQDVGACMILLSVRVFFKKCPSIVQNFAVFPETM 199

Qy   217 P-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ERGFMCQ 271
         :| ||| |||:|| :||| | | |:
```

```
Query Match     46.3%;  Score 679;  DB 13;  Length 988;
Best Local Similarity   50.4%;  Pred. No. 2.2e-55;
Matches  122;  Conservative  51;  Mismatches  57;  Indels  12;  Gaps  5;

Qy    38 IQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQONNWLRTNWVPRNSAQKIYV 97
         :|||||| :|:|:|||||||| |:|| | ||| | |:
Db    20 VTSELAWTTHPEPGWEEVSGYDEAMNPIRTYQVCNVREANONNWLRTKFIQRODVQRVYV 79

Qy    98 ELKFTLRDCNSIPLVLGTCKETFNLYYMESDDD----HGVKFREHQFTKIDTIAADESFT 153
         ||||:|||||| |||||||||||||||||| :|:||||| |:|||||||||
Db    80 ELKFTVRDCNSIPNIPGSCKETFNLFYYESDTDSASANSPFWMENPYIKVDTIAPDESFS 139

Qy   154 QMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFP 213
         :|: |||| |||||||||| :||||||||||||||| |:|||||:|:||||||||
Db   140 KLESG----RVNTKVRSFGPLSKNGFYLAFQDGACMSLISVRAFYKKCSNTIAGFAIFP 195

Qy   214 DTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYEE--RGFM 269
         :|: :| ||| |||:|| :||| | :|
Db   196 EFLTGAEPSLVIAPGTCIPNAVEVSVPLKLYCNGDGEWMVPVGVACTCAAGYEPAMKDTQ 255

Qy   270 CQ 271
         ||
Db   256 CQ 257
```

OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93288394.
RA   SAJJADI F.G., PASQUALE E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
DR   EMBL; Z19110; CAA79526.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
FT   NON_TER   1
FT   NON_TER   1
SQ   SEQUENCE   952 AA; 105994 MW; D939D7D0 CRC32;

Query Match   40.6%; Score 594.5; DB 13; Length 952;
Best Local Similarity   45.3%; Pred. No. 1.8e-47;
Matches 112; Conservative 43; Mismatches 53; Indels 39; Gaps 5;

Qy   32 LLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTIQVCNVMDHSQNNWLRTNWVPRNS   91
               : ||| ||||||: | || ||:||| ||::| ||| ||||| | ::|:|| | : |
Db   3 LMDTRTATAELGWTANPSGWEEVSGYDENLNTIRTYQVCNVFEPNQNNWLLTTFINRRG   62

Qy   92 AQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK-----FREHQFTKIDTIA   147
               :||: ||||:|||||::||:| |:||||||||::| || || ||:|:|:|||
Db   63 AHRIYTEMRFTVRDCSSLPNVPGSCKETFNLYYYETDSVIATKKSAFWTEAPYLKVDTIA   122

Qy   148 ADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVK   207
               ||||:|:| | |::| :|||| |:|: |
Db   123 ADESFSQVDFGGRLMK------------------------------GXFFKKCPSVVQ   150

Qy   208 NLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-   265
               | |:|| | | | ||:|| | | ::|:||:|| ||||||||:|||:|:| |:|||
Db   151 NFAIFPETMTGAESTSLVTARGYCIPNAEEVDVPIKLYCNGDGEWMVPIGRCTCKAGYEP   210

Qy   265 ERGFMCQ   271
               |: |:
Db   211 ENNVACR   217

DR   PFAM; PF01404; EPH_lbd; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF00069; pkinase; 1.
FT   NON_TER   1
SQ   SEQUENCE   938 AA; 103996 MW; 2B01D2A0 CRC32;

Query Match   40.2%; Score 589; DB 11; Length 938;
Best Local Similarity   49.8%; Pred. No. 5.8e-47;
Matches 108; Conservative 45; Mismatches 54; Indels 10; Gaps 4;

Qy   54 EISGVDEHYTPIRTIQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVL   113
               |:||:| |||| |||||| : | ||||| : :||||||||:|||||||:| :
Db   1 EVSGYDEAMPIRIYQVCNVNRESSQNNMLRTGFIWRREVQRYVELKFTVRDCNSIPNIP   60

Qy   114 GTCKETFNLYYMESDDD-----HGVKFREHQFTKIDTIIAADESFTQMDLGDRILKLNTEIR   169
               |:|||||||:|:||| : :|||:|:|:|||| ||||:|||||:| |::||||:
Db   61 GSCKETFNLFYYEADSDVASASSPFWMENPYVKVDTISPDESFSRSMPG----RVNTKVR   116

Qy   170 EVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRG   228
               . ||:.|.||||.||| ||| :|:||| |||| :| ||: | |:|| | | |:|
Db   117 SFGPLSKAGFYLGFQDQGACMSLISVRAFYKKCASTTAGFALPETLLRGAEARSLVIAPR   176

Qy   229 SCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE   264
               :|: :: : | : ||| | :|| |:|| |||:| :
Db   177 ACIANAVEVSVPLKLYCNGDGEWMVPVGACTCATGHE   213

RESULT 14
ID   Q60669   PRELIMINARY;   PRT;   938 AA.
AC   Q60669;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTEIN KINASE (FRAGMENT).
GN   EPHB3 OR ETK2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 9520798.
RA   RUIZ J.C., CONLON F.L., ROBERTSON E.J.;
RT   "Identification of novel protein kinases expressed in the developing mouse heart.";
RL   Mech. Dev. 48:155-164(1994).
DR   EMBL; U11493; AAA67925.1; -.
DR   MGD; MGI:104770; Ephb3.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

RESULT 15
ID   Q91736   PRELIMINARY;   PRT;   902 AA.
AC   Q91736;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   EPH RECEPTOR TYROSINE KINASE (XELK) (FRAGMENT).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96068901.
RA   SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;
RT   "Novel members of the eph receptor tyrosine kinase subfamily expressed
RT   during Xenopus development.";
RL   Oncogene 11:1745-1752(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY (BY
CC   SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC   TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO IN THE BRAIN AND
CC   SPINAL CORD AND IN THE FIRST AND FOURTH VISCERAL ARCHES. MOST
CC   ABUNDANT IN ADULT BRAIN, WITH LOWER LEVELS IN EYE, HEART, OVARY,
CC   OVIDUCT, LUNG AND PHARYNX.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
DR   EMBL; L43621; AAA93527.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;

```
KW   Receptor; Transmembrane; Glycoprotein; Immunoglobulin domain.
FT   NON_TER         1      1
FT   DOMAIN         <1    459      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      460    480      POTENTIAL.
FT   DOMAIN        481    902      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         <1     21      IG-LIKE C2-TYPE DOMAIN (BY SIMILARITY).
FT   DOMAIN        101    237      CYS-RICH.
FT   DOMAIN        238    347      FIBRONECTIN TYPE-III.
FT   DOMAIN        348    445      FIBRONECTIN TYPE-III.
FT   DOMAIN        537    800      PROTEIN KINASE.
FT   DOMAIN        543    551      ATP (BY SIMILARITY).
FT   NP_BIND       543    551      ATP (BY SIMILARITY).
FT   DISULFID        ?     14      BY SIMILARITY.
FT   BINDING       569    569      ATP (BY SIMILARITY).
FT   ACT_SITE      662    662      BY SIMILARITY.
FT   MOD_RES       696    696      PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD      252    252      POTENTIAL.
FT   CARBOHYD      344    344      POTENTIAL.
FT   CARBOHYD      398    398      POTENTIAL.
FT   CARBOHYD      398    398      POTENTIAL.
SQ   SEQUENCE      902 AA;  100850 MW;  C096D006 CRC32;

Query Match             38.4%;  Score 563.5;  DB 13;  Length 902;
Best Local Similarity   55.1%;  Pred. No. 1.4e-44;
Matches  102;  Conservative  42;  Mismatches  34;  Indels   7;  Gaps   4;

Qy    94  KIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVK-----FREHQFTKIDTIAAD 149
          :::||:::||:::||:|.||||||||||||||||:|::  : :  : |:||||||||
Db     2  RVVVEMRFTVRDCSSLPNVPGSCKETFNLYYIEDSNIDNKISTFWNESPYLKVDTIAAD  61

Qy   150  ESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSRVRVIFKKCPFTVKNL 209
          :||:|:| |:|||:| | |.::|:|:   |||||:|||:::|||||:|||| |:|
Db    62  ESFSQVDFGGRLMKVNTEVRSFGPLTRSGFYLAFQDYGACMSLLSVRVFFKKCPSVVQNF 121

Qy   210  AMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPP-RMYCSTEGEWLVPIGKCSCNAGYE-ER 266
          :||:|:| : :|||| ::|::|:| | |::||: :||||||:||||||| :::::| ||||
Db   122  AVPPETMTGAESISLVIARGTCIPNAEEVDVPIKLYCNGDGEWMVPIGKCTCKAGYEPEN 181

Qy   267  GFMCQ 271
          :.|:
Db   182  HVVCK 186
```

Search completed: May 9, 2000, 22:31:15
Job time: 2380 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 22:32:11 ; Search time 49.21 Seconds
(without alignments)
167.716 Million cell updates/sec

Title:
Perfect score: 1466
Sequence: 1 MDCQLSILLLSCSVLDSFG...........VPIGKCSCNAGYEERGFWCQ 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1466	100.0	983	1	EPA3_HUMAN	P29320 homo sapien
2	1400.5	95.5	983	1	EPA3_MOUSE	P29319 mus musculu
3	1399	95.4	984	1	EPA3_RAT	O08680 rattus norv
4	1309	89.3	983	1	EPA3_CHICK	P29318 gallus gall
5	1003	68.4	948	1	EPA6_RAT	P54758 rattus norv
6	1002	68.3	1035	1	EPA6_MOUSE	O62413 mus musculu
7	998	68.1	1035	1	EPA4_CHICK	Q07496 gallus gall
8	995	67.9	986	1	EPA4_HUMAN	P54764 homo sapien
9	995	67.9	986	1	EPA5_CHICK	Q06629 gallus gall
10	990	67.5	877	1	EPA5_MOUSE	Q60629 mus musculu
11	985	67.2	1005	1	EPA5_RAT	P54757 rattus norv
12	984	67.1	1037	1	EPA5_HUMAN	P54756 homo sapien
13	980	66.8	986	1	EPA4_MOUSE	Q03137 mus musculu
14	947	64.6	1013	1	EPA5_CHICK	Q04755 gallus gall
15	927.5	63.3	998	1	EPA7_HUMAN	Q15375 homo sapien
16	925.5	63.2	998	1	EPA7_MOUSE	Q61772 mus musculu
17	925.5	63.1	998	1	EPA7_RAT	P54759 rattus norv
18	925	63.1	1004	1	EPA8_MOUSE	O09127 mus musculu
19	772	52.7	988	1	EPB2_CHICK	P28693 gallus gall
20	767	52.3	986	1	EPA2_HUMAN	P29323 homo sapien
21	767	52.3	993	1	EPB2_MOUSE	P54763 mus musculu
22	760.5	51.9	984	1	EPB1_RAT	P09759 rattus norv
23	752.5	51.3	984	1	EPB1_HUMAN	P54762 homo sapien
24	740	50.5	987	1	EPB2_COTJA	Q90344 coturnix co
25	698.5	47.6	977	1	EPA2_MOUSE	Q03145 mus musculu
26	692	47.2	998	1	EPB3_HUMAN	P54753 homo sapien
27	681	46.5	993	1	EPB3_MOUSE	P54754 mus musculu
28	678	46.2	976	1	EPA2_HUMAN	P29317 homo sapien
29	623	42.5	1002	1	EPB5_CHICK	Q07497 gallus gall
30	604	41.2	976	1	EPA1_HUMAN	P21709 homo sapien
31	528.5	36.1	1006	1	EPB5_HUMAN	O15197 homo sapien
32	510	34.8	987	1	EPB4_MOUSE	P54761 mus musculu
33	492	33.6	987	1	EPB4_HUMAN	P54760 homo sapien
34	99.5	6.8	982	1	MSHM_SARGL	O63852 sarcophyton
35	90	6.1	3110	1	LMA2_HUMAN	P24043 homo sapien
36	88	6.0	1376	1	VGL2_CVM4	P22432 murine coro
37	85	5.8	3106	1	LMA2_MOUSE	Q60675 mus musculu
38	84.5	5.8	380	1	I132_HUMAN	Q14627 homo sapien
39	84	5.7	788	1	ITB3_HUMAN	P05106 homo sapien
40	82.5	5.6	461	1	HMCS_ARATH	P54873 arabidopsis
41	82.5	5.6	1098	1	TSL1_YEAST	P38427 saccharomyc
42	82	5.6	1029	1	YFC5_YEAST	P43571 saccharomyc
43	82	5.6	1376	1	VGL2_CVMJC	O02385 murine coro
44	80.5	5.5	406	1	NPRA_BACST	P43130 bacillus st
45	80.5	5.5	897	1	SAP1_YEAST	P39955 saccharomyc
	80	5.5	806	1	SYL_HELPJ	Q9ZJ63 helicobacte

ALIGNMENTS

RESULT 1
EPA3_HUMAN

ID   EPA3_HUMAN   STANDARD;   PRT;   983 AA.
AC   P29320;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (HEK).
DE   EPHA3 OR ETK1 OR ETK OR HEK.
GN   EPHA3 OR ETK1 OR ETK OR HEK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92147681.
RA   Boyd A.W., Ward L.D., Wicks I.P., Simpson R.J., Salvaris E., Wilks A.,
RA   Welch K., Loudovaris M., Rockman S., Busmanis I.;
RT   "Isolation and characterization of a novel receptor-type protein
RT   tyrosine kinase (hek) from a human pre-B cell line.";
RL   J. Biol. Chem. 267:3262-3267(1992).
RN   [2]
RP   SEQUENCE OF 21-29 AND 840-860, AND CHARACTERIZATION.
RX   MEDLINE; 92179233.
RA   Wicks I.P., Wilkinson D., Salvaris E., Boyd A.W.;
RT   "Molecular cloning of HEK, the gene encoding a receptor tyrosine
RT   kinase expressed by human lymphoid tumor cell lines.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1611-1615(1992).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A2, -A3, -A4 AND -A5. COULD PLAY A ROLE IN LYMPHOID
CC   FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: RESTRICTED TO LYMPHOID TUMOR CELL LINES.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M83941; AAA58633.1; -.
DR   EMBL; A28003; CAA01906.1; -.
DR   PIR; A38224; A38224.
DR   HSSP; P00523; 2PTK.
DR   MIM; 179611;
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00107; PROTEIN_KINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    542    565       POTENTIAL.
FT   DOMAIN      566    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    320       CYS-RICH.
FT   DOMAIN      321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN      432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       PHOSPHORYLATION.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    232    232       POTENTIAL.
FT   CARBOHYD    337    337       POTENTIAL.
FT   CARBOHYD    391    391       POTENTIAL.
FT   CARBOHYD    404    404       POTENTIAL.
FT   CARBOHYD    493    493       POTENTIAL.
FT   CONFLICT    507    507       F -> L (IN CAA01906).
FT   CONFLICT    724    724       V -> L (IN CAA01906).
SQ   SEQUENCE   983 AA;  110086 MW;  B8D900FA80FF5121 CRC64;

Query Match             100.0%;         Score 1466;  DB 1;  Length 983;
Best Local Similarity   100.0%;         Pred. No. 3.1e-122;
Matches  271; Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDE  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDE  60

Qy      61  HYTPIRTYQVCNVMDHSQONNWLRTNWPRNSAQKIYELKFTLRDCNSIPLVLGTCKETF  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  HYTPIRTYQVCNVMDHSQONNWLRTNWPRNSAQKIYELKFTLRDCNSIPLVLGTCKETF  120

Qy     121  NLIYMESDDDHGVKFREHQFTKIDTIAADESETTQMDLGDRILKLNTEIREVGPVNKGFY  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  NLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKGFY  180

Qy     181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240

Qy     241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
            |||||||||||||||||||||||||||||||
Db     241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271

RESULT   2
ID   EPA3_MOUSE      STANDARD;      PRT;    983 AA.
AC   P29319;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1) (MEK4).
GN   EPHA3 OR ETK1 OR MEK4 OR TYRO4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR X SWISS WEBSTER; TISSUE=EMBRYO;
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SECRETED VARIANT WITHOUT THE TRANSMEMBRANE
CC       REGION AND THE KINASE DOMAIN IS PRODUCED BY ALTERNATIVE SPLICING
CC       OF THE SAME GENE.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURRING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M68513; AAA39521.1; -.
DR   EMBL; M68515; AAA39522.1; ALT_SEQ.
DR   PIR; A45583; A45583.
DR   HSSP; P16109; 1FSB.
DR   MGD; MGI:99612; EPHA3.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; Pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    983       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    541    564       POTENTIAL.
FT   DOMAIN      565    983       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    320       CYS-RICH.
FT   DOMAIN      321    431       FIBRONECTIN TYPE-III.
FT   DOMAIN      432    528       FIBRONECTIN TYPE-III.
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       PHOSPHORYLATION.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    231    231       POTENTIAL.
FT   CARBOHYD    336    336       POTENTIAL.
FT   CARBOHYD    390    390       POTENTIAL.
FT   CARBOHYD    403    403       POTENTIAL.
FT   CARBOHYD    492    492       POTENTIAL.
FT   VARSPLIC    530    983       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   983 AA;  109955 MW;  BE44A6655D8107A2 CRC64;

Query Match             95.5%;          Score 1400.5;  DB 1;  Length 983;
Best Local Similarity   95.9%;          Pred. No. 2e-116;
Matches  260; Conservative   3;   Mismatches   7;   Indels   1;   Gaps   1;
```

RESULT 3

```
ID   EPA3_RAT      STANDARD;      PRT;    984 AA.
AC   O08680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR REK4).
GN   EPHA3 OR REK4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE; 98120505.
RA   Li Y.Y., McTiernan C.F., Feldman A.M.;
RT   "IL-1 beta alters the expression of the receptor tyrosine kinase gene
RT   r-EphA3 in neonatal rat cardiomyocytes.";
RL   Am. J. Physiol. 274:H331-H341(1998).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; U69278; AAC06273.1; -.
DR   HSSP; P16105; 1FSB.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00014; IFSB.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
```

```
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    984       EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN       21    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    542    565       POTENTIAL.
FT   DOMAIN      566    984       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       21    321       CYS-RICH.
FT   DOMAIN      322    432       FIBRONECTIN TYPE-III.
FT   DOMAIN      433    529       FIBRONECTIN TYPE-III.
FT   DOMAIN      622    883       PROTEIN KINASE.
FT   NP_BIND     628    636       ATP (BY SIMILARITY).
FT   BINDING     654    654       ATP (BY SIMILARITY).
FT   ACT_SITE    747    747       BY SIMILARITY.
FT   MOD_RES     780    780       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    232    232       POTENTIAL.
FT   CARBOHYD    337    337       POTENTIAL.
FT   CARBOHYD    391    391       POTENTIAL.
FT   CARBOHYD    404    404       POTENTIAL.
FT   CARBOHYD    493    493       POTENTIAL.
SQ   SEQUENCE   984 AA;  110227 MW;  F170C49312F7A0AB CRC64;
```

```
Query Match             95.4%;  Score 1399;  DB 1;  Length 984;
Best Local Similarity   95.2%;  Pred. No. 2.7e-116;
Matches  258;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy    1  MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKRIIQGELGWISYPSHGWEEISGVDE  60
            ||| :|||:| | ||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1  MDCHLSILLFGCCVLSCSRELSPQPSNEVNLLDSKRIIQGELGWISYPSHGWEEISGVDE  60

Qy   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  HYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF  120

Qy  121  NLIYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY  180
            |||||||:|||||:||||||||||||||||||||||||||||||||||||||||||||
Db  121  NLIYMESDDHGVKFLEHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY  180

Qy  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db  181  LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP  240

Qy  241  RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  271
            |||||||||||||||||:|||||||||||::|
Db  241  RMYCSTEGEWLVPIGKCTCNAGYEERGFICQ  271
```

RESULT 4

```
ID   EPA3_CHICK    STANDARD;      PRT;    983 AA.
AC   P29318;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 3 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR ETK1 OR CEK4).
GN   EPHA3 OR ETK1 OR CEK4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92031278.
RA   Sajjadi F.G., Pasquale E.B., Subramani S.;
RT   "Identification of a new eph-related receptor tyrosine kinase gene
RT   from mouse and chicken that is developmentally regulated and encodes
RT   at least two forms of the receptor.";
RL   New Biol. 3:769-778(1991).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
```

```
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: GREATEST LEVELS OF EXPRESSION OCCURRING IN THE
CC       BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; M68514; AAA48666.1; -.
DR   PIR; B45583; B45583.
DR   HSSP; P05523; 2PTK.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT   SIGNAL          1     19      BY SIMILARITY.
FT   CHAIN          20    983      EPHRIN TYPE-A RECEPTOR 3.
FT   DOMAIN         20    540      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      541    564      POTENTIAL.
FT   DOMAIN        565    983      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         20    320      CYS-RICH.
FT   DOMAIN        321    431      FIBRONECTIN TYPE-III.
FT   DOMAIN        432    528      FIBRONECTIN TYPE-III.
FT   DOMAIN        621    882      PROTEIN KINASE.
FT   NP_BIND       627    635      ATP (BY SIMILARITY).
FT   BINDING       653    653      ATP (BY SIMILARITY).
FT   ACT_SITE      746    746      BY SIMILARITY.
FT   MOD_RES       779    779      PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD      231    231      POTENTIAL.
FT   CARBOHYD      336    336      POTENTIAL.
FT   CARBOHYD      390    390      POTENTIAL.
FT   CARBOHYD      403    403      POTENTIAL.
FT   CARBOHYD      492    492      POTENTIAL.
FT   SEQUENCE      983 AA;  109910 MW;  E8895F0BDF77651E CRC64;

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||| ||||||||||| ||
Db   248 EWLVPIGKCLCNAGYEERGFACQ 270

Query Match             89.3%;  Score 1309;  DB 1;  Length 983;
Best Local Similarity   91.6%;
  Matches  241;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     9 LLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRY 68
         ||||  |||||    |||||||||||||||||||||||||||||||||||||||||||
Db     8 LLLLCAALGSAGRLSARPGNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRY 67

Qy    69 QVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFLRDCNSIPLVLGTCKETFNLYYMESD 128
         | |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    68 QESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESD 127

Qy   129 DDHGVKREHQFKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGA 188
         |||   |||||||||||||||||||||||||||||||| |||||| |||||||||||||
Db   128 DDHLAKFRHQFKIDTIAADESFTQMDLGDRILKLNTEVREVGPVSKKGFYLAFQDVGA 187

Qy   189 CVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         ||||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||
Db   188 CVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNHSKEEPPKMYCSTEG 247
```

```
RESULT   5

EPA6_RAT
ID   EPA6_RAT                STANDARD;      PRT;   948 AA.
AC   P54758;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 6 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-2) (EPH HOMOLOGY KINASE-2).
GN   EPHA6 OR EHK2 OR EHK-2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 94067777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
     HSSP; P05523; 2PTK.
     PRINTS; PR00014; FNTYPEIII.
     PRINTS; PR00109; TYRKINASE.
     PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
     PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
     PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
     PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
     PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
     PROSITE; PS01186; EGF_2; UNKNOWN_1.
     PFAM; PF00041; fn3; 2.
     PFAM; PF00069; pkinase; 1.
     PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmembrane; Glycoprotein; Signal.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
FT   SIGNAL          1     22      POTENTIAL.
FT   CHAIN          23    948      EPHRIN TYPE-A RECEPTOR 6.
FT   DOMAIN         23    549      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      550    570      POTENTIAL.
FT   DOMAIN        571    948      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        630    943      PROTEIN KINASE.
FT   NP_BIND       636    644      ATP (BY SIMILARITY).
FT   BINDING       662    662      ATP (BY SIMILARITY).
FT   ACT_SITE      797    797      BY SIMILARITY.
FT   CARBOHYD      342    342      POTENTIAL.
FT   CARBOHYD      396    396      POTENTIAL.
FT   CARBOHYD      409    409      POTENTIAL.
FT   SEQUENCE      948 AA;  106235 MW;  A47DC78EDB2DEF30 CRC64;

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||| ||||||||||| ||
Db   248 EWLVPIGKCLCNAGYEERGFACQ 270

Query Match             68.4%;  Score 1003;  DB 1;  Length 948;
Best Local Similarity   65.7%;
  Matches  180;  Conservative   39;  Mismatches   37;  Indels   18;  Gaps    3;

Qy     3 CQLSILLLSCSVLDSFGELIP----------QPSNEVNLLDSKTIQGELGWISYPSHGW 52
         | | ||| |||||| |||:||          |||:|||||| :| | |||||||| :||
Db     4 CEVREFLL-------QFGFLPLLTAWTGDCSHVSNQVLLDISTVMGELGWISYPLNGW 56

Qy    53 EEISGVDEHYTPIRYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFLRDCNSIPLV 112
         :  : :||:  | ||||||||:  :|||||||||:  : |||||:||||||||||| |
Db    57 DAITEMDEHNRPIHYQVCNVMEPNQNNWLRTNWISRDAAQKIVEMKFTLRDCNSIPWV 116
```

```
Qy    113  LGTCKETFNLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVG  172
           ||||||||:||||:: |||||| || :|||||||||||||||||:||||||:|||||| |||
Db    117  LGTCKETFLTLYIESDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREVG  176

Qy    173  PVNKKGFYLAFQDVGACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCV  231
           |:||||||||||| ||:| | ||||||||||||| ||||||||| | |||:|||||||||||
Db    177  PIERKGFYLAFQDIGACIALVSVSRVYFKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCV  236

Qy    232  NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE  265
           :|||: ||||||||||||||||||| ||| ||||
Db    237  KSSEERDTPKLYCGADGDWLVPLGRCICTGYEE  270
```

RESULT 6

```
ID   EPA6_MOUSE     STANDARD;      PRT;    1035 AA.
AC   Q62413;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 6 PRECURSOR (EC 2.7.1.112) (EPH HOMOLOGY
DE   KINASE RECEPTOR EHK-2) (EPH HOMOLOGY KINASE-2) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-2).
GN   EPHA6 OR EHK2 OR EHK-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE; 97047913.
RA   Lee A.M., Navaratnam D., Ichimiya S., Greene M.I., Davis J.G.;
RT   "Cloning of m-ehk2 from the murine inner ear, an eph family receptor
RT   tyrosine kinase expressed in the developing and adult cochlea.";
RL   DNA Cell Biol. 15:817-825(1996).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; U58332; AAB53836.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:108034; EPHA6.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; Eph_lbd; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Transferase; Tyrosine-protein kinase; Atp-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL          1     22       POTENTIAL.
FT   CHAIN          23   1035       EPHRIN TYPE-A RECEPTOR 6
FT   DOMAIN         23    549       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      550    570       POTENTIAL.
FT   DOMAIN        571   1035       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        630    943       PROTEIN KINASE.
FT   NP_BIND       636    644       ATP (BY SIMILARITY).
FT   BINDING       662    662       ATP (BY SIMILARITY).
FT   ACT_SITE      797    797       BY SIMILARITY.
FT   CARBOHYD      342    342       POTENTIAL.
FT   CARBOHYD      396    396       POTENTIAL.
FT   CARBOHYD      409    409       POTENTIAL.
SQ   SEQUENCE   1035 AA;  116137 MW;  560B26419A5AEF74 CRC64;
```

```
Query Match      68.3%;  Score 1002;  DB 1;  Length 1035;
Best Local Similarity   65.7%;  Pred. No. 4.6e-81;
Matches  180;  Conservative  41;  Mismatches  35;  Indels  18;  Gaps  3;

Qy      3  COLSILILLLSCSVLLDSFGELIP----------QPSNEVNLLDSKTIQGELGWISYPSHGW   52
           : |: |||| ||  |:|   |            || ||| |:| ||    ||| | | |||
Db      4  CEVREFFLL----------QFGFFLPLLTAWTGDCSHVSNQVVLLDTTTVMGELGWKTYPLNGW   56

Qy     53  EEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLV  112
           : |:||:|:|| | ||||||||::|| |||||||||| :|:||||||:|||||||||| |
Db     57  DAITEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIYVEMKFTLRDCNSIPWV  116

Qy    113  LGTCKETFNLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVG  172
           ||||||||:||||:: |||||| || :|||||||||||||||||:||||||:|||||| |||
Db    117  LGTCKETFLTLYIESDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREVG  176

Qy    173  PVNKKGFYLAFQDVGACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCV  231
           |:||||||||||| ||:| | ||||||||||||| ||||||||| | |||:|||||||||||
Db    177  PIERKGFYLAFQDIGACIALVSVSRVYFKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCV  236

Qy    232  NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE  265
           :|||: ||||||||||||||||||| ||| ||||
Db    237  KSSEERDTPKLYCGADGDWLVPLGRCICTGYEE  270
```

RESULT 7

```
ID   EPA4_CHICK     STANDARD;      PRT;    986 AA.
AC   Q074496; Q90772;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR CEK8).
GN   EPHA4 OR CEK8.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPINAL CORD;
RX   MEDLINE; 96404128.
RA   Ohta K., Nakamura M., Hirokawa K., Tanaka S., Iwama A., Suda T.,
RA   Ando M., Tanaka H.;
RT   "The receptor tyrosine kinase, Cek8, is transiently expressed on
RT   subtypes of motoneurons in the spinal cord during development.";
RL   Mech. Dev. 54:59-69(1996).
RN   [2]
RP   SEQUENCE OF 138-986 FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE; 93288294.
RA   Sajjadi F.G., Pasquale E.B.;
RT   "Five novel avian Eph-related tyrosine kinases are differentially
RT   expressed.";
RL   Oncogene 8:1807-1813(1993).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
```

RESULT 8

```
EPA4_HUMAN
ID   EPA4_HUMAN     STANDARD;      PRT;   986 AA.
AC   P54764;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (RECEPTOR PROTEIN-TYROSINE KINASE HEK8)
DE   (EPHA4 OR SEK OR HEK8).
GN   EPHA4 OR SEK OR HEK8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; L36645; AAA74246.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602188; -.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CARBOHYD    545    545       POTENTIAL.
SQ   SEQUENCE    986 AA;  109859 MW;  0C39C1152EDD46F CRC64;
```

```
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
```

```
DR   EMBL; D83174; BAA07373.1; -.
DR   EMBL; Z19059; CAA79509.1; -.
DR   HSSP; P00523; 2PTK.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    986       EPHRIN TYPE-A RECEPTOR 4.
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
FT   NP_BIND     627    635       ATP (BY SIMILARITY).
FT   BINDING     653    653       ATP (BY SIMILARITY).
FT   ACT_SITE    746    746       BY SIMILARITY.
FT   MOD_RES     779    779       PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    340    340       POTENTIAL.
FT   CARBOHYD    408    408       POTENTIAL.
FT   CONFLICT    138    138       S -> R (IN REF. 2).
FT   CONFLICT    487    487       S -> T (IN REF. 2).
SQ   SEQUENCE    986 AA;  109482 MW;  BD88C2A5BD840A0F CRC64;
```

```
Query Match             68.1%;  Score 998;  DB 1;  Length 986;
Best Local Similarity   72.2%;
Matches  179;  Conservative  35;  Mismatches  32;  Indels   2;  Gaps   2;

Qy    26  PSNEVNLLDSKIIQGELGWISYP-SHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRT   84
          :|:||.||||||||::|||:|  | ||||||||||||||||||||:||||||||||||
Db    27  PANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRIYQVCNVMEPSQNNWLRT   86

Qy    85  NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKID  144
          :|:|| .|||:|:|:||||||||:|.|:.||||||||||| ::.|.:.||.|.||:||
Db    87  DWIPREGAQRVYIEIKFTLRDCNSLPGVMGTCKETFNLYYESNNDKERFIRESQFAKID  146

Qy   145  TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSRVYFKKCPF  204
          ||||||||||:|:||||:||||:|.|||:.|||||||||||||||:||||||||:.||
Db   147  TIAADESFTQVDIGDRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSRVYFYKCPL  206

Qy   205  TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGY  263
          ||:||| |||||  .| :|||||||||||||:|:|.|:|||. :|||||||||:|.|||
Db   207  TVRNLAQFPDTITGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGY  266

Qy   264  EERGFMCQ  271
          |||.| ||
Db   267  EERNGECQ  274
```

Query Match   67.9%; Score 995; DB 1; Length 986;
Best Local Similarity   68.8%; Pred. No. 1.8e-80;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

```
Qy     10 LLSC--SVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
          | ||    |||||||||||||||||||||||||||:||   ||:| :|| ||||  ||||
Db      9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy     67 TYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
          |||||||||||||||||||||||||||||||||||:||:|| |||:||| ||  ||||||
Db      9 TYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVWGTCKETFNLYYE 68

Qy    127 SDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
          :|||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db    129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKRKGFYILAFQDV 188

Qy    187 GACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS 245
          |||:||||||||||||||:|||||||||||| ||:|||| ||||||||||||::|:||
Db    189 GACIALVSVSRVYFKKCPLTVRNLAQPFDTITGADSSLVEVRGSCVNNSEEKDVPRKMYCG 248

Qy    246 TEGEWLVPIGKCSCNAGYEERGFMCQ 271
          :|||||||||| |||||| || ||
Db    249 ADGEWLVPIGNCLCNAGHEERSGECQ 274
```

RESULT   9
ID   EPA5_MOUSE   STANDARD;   PRT;   877 AA.
AC   Q60629;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (BRAIN-SPECIFIC KINASE)
DE   (CEK-7).
DE   EPHA5 OR EHK1 OR CEK7 OR BSK.
GN   EPHA5 OR EHK1 OR CEK7 OR BSK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 94194581.
RA   Zhou R., Copeland T.D., Kromer L.F., Schulz N.T.;
RT   "Isolation and characterization of Bsk, a growth factor receptor-like
RT   tyrosine kinase associated with the limbic system.";
RL   J. Neurosci. Res. 37:129-143(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; U07357; AAA17038.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:99654; EPHA5.
DR   PRINTS; PR00014; FNTPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
PFAM; PF00041; fn3; 1.
PFAM; PF00069; pkinase; 1.
PFAM; PF00536; SAM; 1.
PFAM; PF01404; EPH_lbd; 1.
KW   Receptor; Transmerase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Transferase; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL   1   26   BY SIMILARITY.
FT   CHAIN   27   877   EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN   27   412   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   413   433   POTENTIAL.
FT   DOMAIN   434   877   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   305   399   FIBRONECTIN TYPE-III.
FT   DOMAIN   514   775   PROTEIN KINASE.
FT   NP_BIND   520   528   ATP (BY SIMILARITY).
FT   BINDING   546   546   ATP (BY SIMILARITY).
FT   ACT_SITE   639   639   BY SIMILARITY.
FT   CARBOHYD   266   266   POTENTIAL.
FT   CARBOHYD   301   301   POTENTIAL.
SQ   SEQUENCE   877 AA; 97115 MW; 54AD2DC86417821A CRC64;

```
Qy     10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
          || ||: | || |:|  ||||||||||:||:|| || |  ||||||| ||  |||||
Db     46 LLLCAALRT---LLASPSNEVNLLDSRTVMGDLGWIAPFKNGWEEIGEVDENYAPIHTYQ 102

Qy     70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
          |||  |:||||:|::| :||| |  | ||:|||||||||:| |||:||| ||  :|||
Db    103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNLYYFESDD 162

Qy    130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
          :||||||||||||||||||||||| ||| :| |::|| |||| | | |||||:|||||
Db    163 ENGRSIKENQYIKIDTIAADESFTELDLGDRVWKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy    190 VALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
          :|||||||||||| :||:|| |:||| | ||:|| ||||| | ||| ::| | || |
Db    223 IALVSVSRVYIKKCPSVVRHLAIIPDTITGADSSLQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy    249 EWLVPIGKCSCNAGYEERGFMCQ 271
          |||||||| || ||| |||||||
Db    283 EWLVPIGKCNCKAGYEEKNGTCQ 305
```

RESULT   10
ID   EPA5_RAT   STANDARD;   PRT;   1005 AA.
AC   P54757;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1).
GN   EPHA5 OR EHK1 OR EHK-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX   MEDLINE; 94067777.
RA   Maisonpierre P.C., Barrezueta N.X., Yancopoulos G.D.;
RT   "Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine
RT   kinase family with distinctive structures and neuronal expression.";
RL   Oncogene 8:3277-3288(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;

Query Match   67.5%; Score 990; DB 1; Length 877;
Best Local Similarity   66.2%; Pred. No. 4.3e-80;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
RX   MEDLINE; 95206467.
RA   Taylor V., Pfarr S., Miescher G.C., Honegger P., Breitschopf H.,
RA   Lassmann H., Steck A.J.;
RT   "Expression and developmental regulation of Ehk-1, a neuronal
RT   Elk-like receptor tyrosine kinase in brain.";
RL   Neuroscience 63:163-178(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   -!- EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   -!- PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC   -!- ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC   -!- SYSTEM. PREDOMINANTLY EXPRESSED IN NEURONS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   -!- DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; X78689; CAA55357.1; -.
DR   HSSP; P00523; 2PTK.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL         1     ?
FT   CHAIN          ?    1005       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN         ?     575       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      576    596       POTENTIAL.
FT   DOMAIN        597   1005       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        677    938       PROTEIN KINASE.
FT   NP_BIND       683    691       ATP (BY SIMILARITY).
FT   BINDING       709    709       ATP (BY SIMILARITY).
FT   ACT_SITE      802    802       BY SIMILARITY.
FT   CARBOHYD      266    266       POTENTIAL.
FT   CARBOHYD      301    301       POTENTIAL.
FT   CARBOHYD      371    371       POTENTIAL.
FT   CARBOHYD      425    425       POTENTIAL.
FT   CARBOHYD      438    438       POTENTIAL.
FT   CARBOHYD      463    463       POTENTIAL.
FT   CARBOHYD       10     20       POTENTIAL.
FT   VARSPLIC      597    621       GRRRTQGRGGG -> DADGPRAQASWCHARR (IN
FT                                  SPLICED FORMS).
FT   VARSPLIC      306    358       VCRPGFKRASPHSQTCSKCPPHSYTHEEASTSCVCEKDYFR
FT                                  RESDPPTWACTR -> G (IN SPLICED FORMS).
FT   VARSPLIC      358    470       RPSAPRNAISNVNETSVFLEWIPPADTGGGKDVSYIICK
FT                                  KNSHAGVCEEGGHVRYLPQOIGLKNTSVMADPLAHTNY
FT                                  TPIEAVNGVSDLSPGTRQYVSVNVTINQAA -> T (IN
FT                                  SPLICED FORMS).
FT   VARSPLIC                       SSSCEGCGCGRASSLCAVAHPSLIW -> R (IN
FT                                  SPLICED FORMS).
FT   CONFLICT      170    170       D -> E (IN REF. 2).
FT   CONFLICT      566    566       G -> A (IN REF. 2).
FT   CONFLICT      578    578       G -> A (IN REF. 2).
FT   CONFLICT      669    669       G -> A (IN REF. 2).
FT   CONFLICT      708    708       T -> I (IN REF. 2).
FT   CONFLICT      979    979       T -> I (IN REF. 2).
```

```
SQ   SEQUENCE   1005 AA;  111007 MW;  1AED42C99693C574 CRC64;

RESULT 11

ID   EPA5_HUMAN     STANDARD;      PRT;   1037 AA.
AC   P54756;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-1) (EPH HOMOLOGY KINASE-1) (RECEPTOR PROTEIN-
DE   TYROSINE KINASE HEK7).
GN   EPHA5 OR EHK1 OR HEK7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Miescher G.C.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 25-1037 FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC   -!- EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC   -!- PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC   -!- ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ALMOST EXCLUSIVELY EXPRESSED IN THE NERVOUS
CC   -!- SYSTEM.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC   -!- DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
Query Match          67.2%;  Score 985;  DB 1;  Length 1005;
Best Local Similarity   65.8%;
Matches  173;  Conservative  45;  Mismatches  41;  Indels  4;  Gaps  2;

Qy    10 LLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHTPIRTYQ 69
         || | |: ||||||| ||||||:||:|:|||||||| |||| ||:||:||| |||
Db    46 LLLCAALRT---LLASPSNEVNLLDSRVIVGLDLGWIAPPKNGWEEIGEVDENYAPIHTYQ 102

Qy    70 VCNVMDHSQNNWLRTNWWPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         ||:||| || ||| ||:||||||||||||||||||||:|:||| ||||||:|||
Db   103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNNYYFESDD 162

Qy   130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         |: : ||:|| ||||||||||||| ||||| ||||||:|||| ||||||||| |||||
Db   163 ENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy   190 VALVSVRYFKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         |||||||| || :| || | ::| |: ||:|||||||||| |:|| ||||||
Db   223 IALVSVRYVKYKKCPSVVRHLAVFPDTIIGADSSQOLLEVSGSCVNHSVTDDPPRKMHCSAEG 282

Qy   249 EWLYPIGKCSCNRAGYEERGFMCQ 271
         ||||||| :|| |||  |  |
Db   283 EWLVPIGKCMCKAGYEEKNGTCQ 305
```

```
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb.sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X95425; CAA64700.1; -.
DR   EMBL; L36644; AAA74245.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 600004; -.
RP   PRINTS; PR00014; FNTYPEIII.
RP   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     24
FT   CHAIN        25   1037       EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN       25    573       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    574    594       POTENTIAL.
FT   DOMAIN      595   1037       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      675    936       PROTEIN KINASE.
FT   NP_BIND     681    689       ATP (BY SIMILARITY).
FT   BINDING     707    707       ATP (BY SIMILARITY).
FT   ACT_SITE    800    800       BY SIMILARITY.
FT   BINDING     707    707       ATP (BY SIMILARITY).
FT   CARBOHYD    264    264       POTENTIAL.
FT   CARBOHYD    299    299       POTENTIAL.
FT   CARBOHYD    369    369       POTENTIAL.
FT   CARBOHYD    423    423       POTENTIAL.
FT   CARBOHYD    436    436       POTENTIAL.
FT   CARBOHYD    461    461       POTENTIAL.
FT   VARSPLIC    597    619       SCCECGCGRASSLCAVAHPILIW -> R (IN ISOFORM
FT                               2).
SQ   SEQUENCE   1037 AA;  114784 MW;  FC2C46C959AFB699 CRC64;

Query Match             67.1%;
Best Local Similarity   65.8%;
Matches   173;  Conservative   44;  Mismatches   42;  Indels    4;  Gaps    2;

Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ 69
         |  :||  ||||:||  :|||::  |||||||:|||||: |||||||||:::||::|||:|||
Db   44 LLLCAALRT---LLASPSNEVNLLDSRTVVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 100

Qy   70 VCNVMDHSQNNWLRTNWPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLIYMESDD 129
         :|::||||||: ||  |  |:|  ||||:||::||||||||| ||:|| ||:|:|||||
Db  101 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMIYFESDD 160

Qy  130 DHGVKFREHQFTKIDTIAADESFTQKDLGIDGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         :|:| ||||||||||:|||||||:||||:|||||||||||:|||||||::|:||||||||||:
Db  161 QNGRNIKENQYIKIDTIAADESFTELDLGDRVWKLNTEVRDVGPLSKKGFYLAFQDVGAC 220

Qy  190 VALVSVRVYIEKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         :||||||||||:|||  ::|  |  || || |  : || ||||||||:|||||||::|| ||
Db  221 IALVSVRVYYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSCVNHSVTDEPPKMHCSAEG 280

Qy  249 EWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||| | |||||||||:|
Db  281 EWLVPIGKCMCKAGYEEKNGTCQ 303

RESULT   12
EPA4_MOUSE
ID   EPA4_MOUSE     STANDARD;      PRT;   986 AA.
AC   Q03137;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
```

```
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 4 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR SEK) (MPK-3).
GN   EPHA4 OR SEK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL; TISSUE=EMBRYONIC BRAIN;
RX   MEDLINE; 93096484.
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattel M.-G.,
RA   Chestier A., Wilkinson D.G., Charnay P.;
RL   Oncogene 8:1103-1103(1993).
RN   [2]
RP   ERRATUM.
RX   MEDLINE; 93205393.
RA   Gilardi-Hebenstreit P., Nieto M.A., Frain M., Mattel M.-G.,
RA   Chestier A., Wilkinson D.G., Charnay P.;
RL   Oncogene 7:2499-2506(1992).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A4 AND -A5. BINDS MORE POORLY TO EPHRIN-A2 AND A-3.
CC       MAY PLAY A ROLE IN A SIGNAL TRANSDUCTION PROCESS INVOLVED IN
CC       HINDBRAIN PATTERN FORMATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A SHORTER FORM WITH A DELETION OF 49
CC       AMINO ACIDS ALTERING THE CATALYTIC SITE MAY BE GENERATED BY
CC       ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN THE BRAIN, LOWER LEVELS
CC       IN THE HEART, LUNG AND KIDNEY. IT IS SEGMENTALLY EXPRESSED
CC       IN THE EMBRYONIC BRAIN.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X65138; CAA46268.1; -.
DR   EMBL; X57241; CAA40517.1; -.
DR   EMBL; S57168; AAB25836.1; -.
DR   HSSP; P00523; 2PTK.
MGD; MGI:98277; EPHA4.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     19
FT   CHAIN        20    986
FT   DOMAIN       20    547       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    548    569       POTENTIAL.
FT   DOMAIN      570    986       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      325    435       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      436    532       FIBRONECTIN TYPE-III (BY SIMILARITY).
FT   DOMAIN      621    882       PROTEIN KINASE.
```

```
FT   NP_BIND     627    635   ATP (BY SIMILARITY).
FT   BINDING     653    653   ATP (BY SIMILARITY).
FT   ACT_SITE    746    746   BY SIMILARITY.
FT   MOD_RES     779    779   PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT   CARBOHYD    235    235   POTENTIAL.
FT   CARBOHYD    340    340   POTENTIAL.
FT   CARBOHYD    408    408   POTENTIAL.
FT   CARBOHYD    423    423   POTENTIAL.
FT   VARSPLIC    783    832   MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE    986 AA;   109801 MW;   D16AD8B5668C80E CRC64;

Query Match               66.8%;   Score 980;   DB 1;   Length 986;
Best Local Similarity     71.4%;   Pred. No. 3.0e-79;
Matches  177;   Conservative  35;   Mismatches  34;   Indels  2;   Gaps  2;

Qy   26   PSNEVNLLDSKIIQGELGLWISYP-SHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRT    84
          |:|||  ||||||:||:: |||| :|  |||||||:||||||||||||||||||| ||||||
Db   27   PANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRIYQVCNVMEASQNNWLRT    86

Qy   85   NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETKETFNLYYMESDDHGVKFREHQFTKID  144
          :|:|| |||||||||||||:||||||||| |||||||| |||||  ::::|:| :|| |:||
Db   87   DWITREGAQRVYIEIKFTLRDCNSLPGVYMGTICKETFNLYYESDNDKRFIRESQFGKTD   146

Qy   145  TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSRVYFKKCPF    204
          ||||||||||:|:||||:||||||| |||:.|||||||||||||||:||||||:.||||:
Db   147  TIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSRVFYKKCPL    206

Qy   205  TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWMLVPIGKCSCNAGY   263
          ||:||| |||:. |||:|||||||||||||:|:.|.|:||| .|||::|||||:| ||||:
Db   207  TVRNLAQFPDTIIGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGH   266

Qy   264  EERGFMCQ  271
          ||: | ||
Db   267  EEONGECQ  274
```

RESULT 13
EPA5_CHICK

```
ID   EPA5_CHICK     STANDARD;      PRT;    1013 AA.
AC   P54755;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR CEK7).
GN   EPHA5 OR CEK7.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC   Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BODY WALL;
RX   MEDLINE; 95047429.
RA   Siever D.A., Verderame M.F.;
RT   "Identification of a complete Cek7 receptor protein tyrosine kinase
RT   coding sequence and cDNAs of alternatively spliced transcripts.";
RL   Gene 148:219-226(1994).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: A NUMBER OF VARIANTS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
```

```
DR   EMBL; U03910; AAB60613.1; -.
DR   EMBL; U03910; AAB60614.1; -.
DR   EMBL; U03910; AAB60612.1; -.
DR   HSSP; P05523; 2PTK.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS50041; fn3; 2.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     31   POTENTIAL.
FT   CHAIN        32   1013   EPHRIN TYPE-A RECEPTOR 5.
FT   DOMAIN        1     31   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    550    570   POTENTIAL.
FT   DOMAIN      571   1013   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      651    912   PROTEIN KINASE.
FT   NP_BIND     657    665   ATP (BY SIMILARITY).
FT   BINDING     683    683   ATP (BY SIMILARITY).
FT   ACT_SITE    776    776   BY SIMILARITY.
FT   CARBOHYD    240    240   POTENTIAL.
FT   CARBOHYD    275    275   POTENTIAL.
FT   CARBOHYD    345    345   POTENTIAL.
FT   CARBOHYD    399    399   POTENTIAL.
FT   CARBOHYD    412    412   POTENTIAL.
FT   CARBOHYD    437    437   POTENTIAL.
FT   VARSPLIC    280    280   MISSING (IN ISOFORM 2).
FT   VARSPLIC    573    595   SCCDHGCGWASSLRAVAYPSLIW -> R (IN ISOFORM
FT                            1 AND ISOFORM 2).
SQ   SEQUENCE   1013 AA;   112245 MW;   AC36FD6AEBF38382 CRC64;

Query Match               64.6%;   Score 947;   DB 1;   Length 1013;
Best Local Similarity     64.6%;   Pred. No. 3.4e-76;
Matches  170;   Conservative  37;   Mismatches  52;   Indels  4;   Gaps  2;

Qy   10   LLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ    69
          ||  : ||:: ||||::|||||||||:|||||||||:|||||||||||||||||||||||
Db   20   LLLCAALRS---LLASPGSEVNLLDSRTVWGDLGWIAYPKNGWEEIGEVDENYAPIHTYQ    76

Qy   70   VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD   129
          ||:||::.|||||.::|: ||.:: .||:|||||||||||:|  ||||||||:||:||||
Db   77   VCKVMEQNQNNWLLTSWISNEGRPASSFELKFTLRDCNSLPGGLGTCKETFNMYYFESDD   136

Qy   130  DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC   189
          : | : ||:|:.|||||||||||||:|||:::.||:||||||||:.|:||||||||||||
Db   137  EDGRNIRENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLTKGFYLAFQDVGAC   196

Qy   190  VALVSVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG   248
          :|||||| |||||.|::|||.||||:. ||:|||||.|.|||||:.:.:::.|.||:||
Db   197  IALVSVRVYKKCPSVIRNLARFPDTIIGADSSQLLEVSGVCVNNHSVTDEAPKMHCSAEG  256

Qy   249  EWLVPIGKCSCNAGYEERGFMCQ   271
          ||||||||| |:||:||:.| ||
Db   257  EWLVPIGKCLCKAGYEEKNNTCQ   279
```

RESULT 14
EPA7_HUMAN

```
ID   EPA7_HUMAN     STANDARD;      PRT;    998 AA.
AC   Q15375;
DT   01-NOV-1997 (Rel. 35, Created)
```

```
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3) (RECEPTOR PROTEIN-
DE   TYROSINE KINASE HEK11).
GN   EPHA7 OR EHK3 OR HEK11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95206782.
RA   Fox G.M., Holst P.L., Chute H.T., Lindberg R.A., Janssen A.M.,
RA   Basu R., Welcher A.A.;
RT   "cDNA cloning and tissue distribution of five human EPH-like receptor
RT   protein-tyrosine kinases.";
RL   Oncogene 10:897-905(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC       --------------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC       --------------------------------------------------------------------------
DR   EMBL; L36642; AAA74243.1; -.
DR   HSSP; P00523; 2PTK.
DR   MIM; 602190; -.
DR   PFAM; PF00041; fn3; 2.
DR   PFAM; PF00069; pkinase; 1.
DR   PFAM; PF00536; SAM; 1.
DR   PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    998       EPHRIN TYPE-A RECEPTOR 7.
FT   DOMAIN       25    556       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    557    577       POTENTIAL.
FT   DOMAIN      578    998       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      329    438       FIBRONECTIN TYPE-III.
FT   DOMAIN      439    537       FIBRONECTIN TYPE-III.
FT   DOMAIN      633    894       PROTEIN KINASE.
FT   NP_BIND     639    647       ATP (BY SIMILARITY).
FT   BINDING     665    665       ATP (BY SIMILARITY).
FT   ACT_SITE    758    758       BY SIMILARITY.
FT   MOD_RES     791    791       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD    343    343       POTENTIAL.
FT   CARBOHYD    410    410       POTENTIAL.
SQ   SEQUENCE   998 AA;  112096 MW;  479B9CA0D2BB06EB CRC64;
Qy     25  QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVADHSQNNWLRT     84
            | ||||| |||||||| |||||||||  ||:|||||:|||||||  ||||:|:||||||||
Db     28  QAAKEVLLLDSKAQQTFLEWISSPPNGWEEISGLDENYTPIRTYQVCQVMEPNQNNWLRT     87
Qy     85  NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLIYMESDDHGVKFREHQFTKID    144
            |::|:||:|:|||||||||||||:||:|||:|||||||:|||:||:| ||:|: ||||
Db     88  NWISKGNAQRIFVELKFTLRDCNSLPLVGTCKETFNLYYYETDIDTGRNIRENLYVKID    147
Qy    145  TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPF    204
            |||||||||:| | ||:||||||:|||||||||||| ||||||||:|||||:||||||||
Db    148  TIAADSSFTQGDLGERKMKLNTEVREIGPLSKKGFYLAFQDVGACIALVSVKVYIKKCWS    207
Qy    205  TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNA    261
            ||||| |||||:|  | ||:|||||:||:|:||  ||||||||| |||||||||||||||
Db    208  IIENLAIFPDDVTGSEFSSLVEVRGTCVSSAEEEAENAPRMHCSAEGEWLVPIGKCICKA    267
Qy    262  GYEERGFMCQ    271
            ||:||||||
Db    268  GYQQKGDTCE    277
```

```
RESULT  15
EPA7_MOUSE
ID   EPA7_MOUSE     STANDARD;      PRT;     998 AA.
AC   Q61772; Q61773; Q61774; Q61505;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   EPHRIN TYPE-A RECEPTOR 7 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR EHK-3) (EPH HOMOLOGY KINASE-3) (EMBRYONIC BRAIN
DE   KINASE) (EBK) (DEVELOPMENTAL KINASE 1) (MDK-1).
GN   EPHA7 OR EHK3 OR EBK OR MDK1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 95124729.
RA   Ciossek T., Millauer B., Ullrich A.;
RT   "Identification of alternatively spliced mRNAs encoding variants of
RT   MDK1, a novel receptor tyrosine kinase expressed in the murine
RT   nervous system.";
RL   Oncogene 10:97-108(1995).
RN   [2]
RP   SEQUENCE OF 431-998 FROM N.A. (ISOFORM MDK1).
RC   TISSUE=BRAIN;
RX   MEDLINE; 96081374.
RA   Ellis J., Liu Q., Breitman M., Jenkins N.A., Gilbert D.J.,
RA   Copeland N.G., Tempest H.V., Warren S., Muir E., Schilling H.,
RA   Fletcher F.A., Ziegler S.F., Rogers J.H.;
RT   "Embryo brain kinase: a novel gene of the eph/elk receptor tyrosine
RT   kinase family.";
RL   Mech. Dev. 52:319-341(1995).
CC   -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC       EPHRIN-A1, -A2, -A3, -A4 AND -A5.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC       PROTEIN TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: FIVE FORMS OF THE PROTEIN (MDK1, MDK1-1,
CC       MDK1-2, MDK1-T1 AND MDK1-T2) ARE PRODUCED BY ALTERNATIVE
CC       SPLICING OF THE SAME GENE. THE SEQUENCE SHOWN HERE IS THAT OF
CC       MDK1.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN EMBRYO. IN ADULT,
CC       EXPRESSION RESTRICTED TO HIPPOCAMPUS, TESTIS AND SPLEEN.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC       --------------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
```

```
Query Match             63.3%;   Score 927.5;  DB 1;  Length 998;
Best Local Similarity   68.4%;   Pred. No. 1.8e-74;
Matches  171; Conservative  39; Mismatches  37; Indels   3; Gaps   2;
```

```
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X79082; CAA55687.1; -.
DR   EMBL; X79083; CAA55688.1; -.
DR   EMBL; X79084; CAA55689.1; -.
DR   EMBL; X81466; CAA57224.1; -.
DR   HSSP; P00523; 2PTK.
DR   MGD; MGI:95276; EPHA7.
DR   PRINTS; PR00014; FNTYPEIII.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR   PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
     PFAM; PF00041; fn3; 2.
     PFAM; PF00069; pkinase; 1.
     PFAM; PF00536; SAM; 1.
     PFAM; PF01404; EPH_lbd; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW   Receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.
FT   SIGNAL        1     29   POTENTIAL.
FT   CHAIN        30    998   EPHRIN TYPE-A RECEPTOR 7.
FT   DOMAIN       30    556   EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      557    577   POTENTIAL.
FT   TRANSMEM    578    998   POTENTIAL.
FT   DOMAIN      329    438   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      439    537   FIBRONECTIN TYPE-III.
FT   DOMAIN      633    894   FIBRONECTIN TYPE-III.
FT   DOMAIN      639    647   PROTEIN KINASE.
FT   NP_BIND     665    665   ATP (BY SIMILARITY).
FT   BINDING     758    758   ATP (BY SIMILARITY).
FT   ACT_SITE    791    791   BY SIMILARITY.
FT   MOD_RES     791    791   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CARBOHYD     64     64   POTENTIAL.
FT   CARBOHYD    343    343   POTENTIAL.
FT   CARBOHYD    410    410   POTENTIAL.
FT   CARBOHYD    544    544   POTENTIAL.
FT   VARSPLIC    604    610   MISSING (IN ISOFORM MDK1-T1).
FT   VARSPLIC    600    610   FKPPGTKTYID -> SLVTNEHLSVL (IN ISOFORM
FT                            MDK1-T1).
FT   VARSPLIC    601    604   MISSING (IN ISOFORM MDK1-2).
FT   VARSPLIC    540    544   MISSING (IN ISOFORM MDK1-1).
FT   VARSPLIC    600    626   EKFQHNKKWMIASCSRL (IN ISOFORM MDK1-T2).
FT   VARSPLIC    611    998   FKFPGTKIYIIDPETYEDPNRAVHQFAK -> SLYRERGDGM
FT                            (IN ISOFORM MDK1-T2).
FT   CONFLICT    480    480   Y -> H (IN REF. 2).
SQ   SEQUENCE    998 AA;  111874 MW;  81C5538E15AEA2FA CRC64;

Query Match      63.28;   Score 926.5;   DB 1;   Length 998;
Best Local Similarity    68.8%;
Matches  172;  Conservative  37;  Pred. No. 2.2e-74;  Mismatches  38;  Indels  3;  Gaps  2;

Qy    25 QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
            | | |||||| ||  |    | :|||||||||||:||||||||||:|| :|||||:|
Db    28 QAAKEVLLLDSKAQQTELEWISSPSGWEEISGLDENYTPIRTYQVCQVMEPNONNWLRT 87

Qy    85 NWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKID 144
            |: :|||  :|| :||||:|||| |||||||||||:|||||||||:|:| | |||||
Db    88 NWISKGNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLIYYETDYDTGRNIRENLYVKID 147

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPF 204
            |||||||| |||| || | | |:  :|  |   || ||||||| |||:|||||||||||
Db   148 TIAADESFTQGDLGERKMKLNTEVREIGPLSKKGFYLAFQDVGACIALVSVKVYYKKCWT 207

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCSCNA 261
            ||||||| |:   | ||| | | :|| | || |     :|||| |||||||||| ||
Db   208 IVENLAVFPDTVTGSEFSSLVEVRGTCVSSAEEAENSPRMHCSAEGEWLVPIGKCICKA 267

Qy   262 GYEERGFMCQ 271
            |:::|| |
Db   268 GYQQKGDTCE 277
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 9, 2000, 22:27:18 ; Search time 64.29 Seconds
                   (without alignments)
                   247.140 Million cell updates/sec

Title: US-09-104-340-4
Perfect score: 1466
Sequence: 1 MDCQLSILLLLSCSVLDSFG.........VPIGKCSCNAGYEERGFMCQ 271

Scoring table: BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 5862943 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                   Listing first 45 summaries

Database :   PIR_63:*
                   1: pir1:*
                   2: pir2:*
                   3: pir3:*
                   4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1466	100.0	983	2	A38224	protein-tyrosine k
2	1400.5	95.5	983	2	A45583	receptor tyrosine
3	1309	89.3	983	2	B45583	receptor tyrosine
4	1003	68.4	948	2	S51605	receptor tyrosine
5	995	67.9	986	2	I78844	receptor protein-t
6	990	67.5	877	2	I48967	brain-specific kin
7	988	67.4	898	2	S47489	receptor protein-t
8	985	67.2	893	2	S51603	receptor-like tyro
9	985	67.2	981	2	S51604	receptor-like tyro
10	985	67.2	1005	2	S49015	receptor-like tyro
11	984	67.1	991	2	I78843	receptor protein-t
12	980	66.8	986	2	S78059	receptor tyrosine
13	977	66.6	985	2	I51149	protein-tyrosine k
14	947	64.6	1013	2	I50615	receptor-type prot
15	927.5	63.3	998	2	I58351	receptor protein-t
16	926.5	63.2	605	2	JC5673	receptor tyrosine
17	926.5	63.2	610	2	I48612	developmental kina
18	926.5	63.2	616	2	I48614	developmental kina
19	926.5	63.2	998	2	JC5672	receptor tyrosine
20	772	52.7	995	2	A56599	embryo kinase 5 -
21	768.5	52.4	985	2	I51672	receptor tyrosine
22	767	52.3	970	2	I78842	receptor protein-t
23	760.5	51.9	984	2	A39753	protein-tyrosine k
24	698.5	47.6	998	2	S49204	tyrosine kinase Mp
25	692	47.2	998	2	S37627	protein-tyrosine k
26	681	46.5	993	2	I48653	mouse developmenta
27	679	46.3	976	2	I50611	protein-tyrosine k
28	678	46.2	976	2	A36355	protein-tyrosine k
29	662	45.2	975	2	I48974	receptor-protein t
30	604	41.2	984	1	A34076	protein-tyrosine k
31	594.5	40.6	952	2	I50612	protein-tyrosine k
32	589	40.2	938	2	I49071	protein kinase - m
33	569.5	38.8	849	2	I50617	protein-tyrosine k
34	528.5	36.1	1006	2	JC5526	kinase-defective E
35	510	34.8	987	2	I48953	eph-related recept
36	493	33.6	987	2	I48652	mouse developmenta
37	397.5	27.1	1019	2	A54092	protein-tyrosine k
38	250.5	17.1	1122	2	T13039	tyrosine kinase re
39	96	6.5	982	2	T42400	tyrosine kinase re
40	93.5	6.4	698	2	H71401	Eph receptor tyros
41	93	6.4	668	1	S58881	probable A6 anther
42	89.5	6.1	982	2	T05803	muts protein homol
43	89	6.0	1376	1	VGIHJ2	hypothetical prote
44	87.5	6.0	788	2	I51530	E2 glycoprotein pr
45	86	5.9	919	2	T29581	Integrin beta-3 su

ALIGNMENTS

RESULT 1
A38224
protein-tyrosine kinase (EC 2.7.1.112) hek precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C;Accession: A38224; B38224
R;Wicks, I.P.; Wilkinson, D.; Salvaris, E.; Boyd, A.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 1611-1615, 1992
A;Title: Molecular cloning of HEK, the gene encoding a receptor tyrosine kinase expre
A;Reference number: A38224; MUID:92179233
A;Accession: A38224
A;Molecule type: mRNA
A;Residues: 1-983 <WIC>
A;cross-references: GB:M83941; NID:g183931; PIDN:AAA58633.1; PID:g183932
A;Experimental source: pre-B-cell leukemia cell line LK63
A;Note: sequence extracted from NCBI backbone (NCBIP:86627)
A;Accession: B38224
A;Molecule type: protein
A;Residues: 21-39;810-860 <WI2>
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferas
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-983/Product: protein-tyrosine kinase hek #status experimental <MAT>
F;542-565/Domain: transmembrane #status predicted <TMM>
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;232,337,391,404,493/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match   100.0%;   Score 1466;   DB 2;   Length 983;
Best Local Similarity   100.0%;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60

Qy   61 HYPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 HYPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120

Qy 121 NLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180

Qy 181 LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240

Qy 241 RMCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
         ||||||||||||||||||||||||||||||
Db 241 RMCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271

RESULT 2

A45583

receptor tyrosine kinase Mek4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse an
A;Reference number: A45583; MUID:92031278
A;Accession: A45583
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68513; NID:g199119; PIDN:AAA39521.1; PID:g199120
A;Note: sequence extracted from NCBI backbone (NCBIN:62398, NCBIP:62401)
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III
C;Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   95.5%;   Score 1400.5;   DB 2;   Length 983;
Best Local Similarity   95.9%;   Pred. No. 1.9e-116;
Matches 260;   Conservative 3;   Mismatches 7;   Indels 1;   Gaps 1;

Qy   1 MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDCHLSLVLLGCCVLSCSGELSPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDE 60

Qy   61 HYPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 HYPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETF 120

Qy   121 NLIYMESDDDHGYKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NLIYMES--DDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFY 179

Qy   181 LAFQDVGACVALVSVRVRYFKKCPPTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 LAFQDVGACVALVSVRVRYFKKCPPTVKNLAMFPDTVPMDSQSLVEVRGSCVNNSKEEDPP 239

Qy   241 RMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
            ||||||||||||||||||||||||||||||||
Db   240 RMYCSTEGEWLVPIGKCTCNAGYEERGFICQ 270

RESULT 3

B45583

receptor tyrosine kinase Cek4 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: B45583
R;Sajjadi, F.G.; Pasquale, E.B.; Subramani, S.
New Biol. 3, 769-778, 1991
A;Title: Identification of a new eph-related receptor tyrosine kinase gene from mouse an
A;Reference number: A45583; MUID:92031278
A;Accession: B45583
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-983 <SAJ>
A;Cross-references: GB:M68514; NID:g454809; PIDN:AAA48666.1; PID:g211447
A;Note: sequence extracted from NCBI backbone (NCBIN:62405, NCBIP:62411)
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; phosphoprotein; phosphoprotein; transmembrane protein
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif

Query Match   89.3%;   Score 1309;   DB 2;   Length 983;
Best Local Similarity   91.6%;   Pred. No. 2.6e-108;
Matches 241;   Conservative 8;   Mismatches 14;   Indels 0;   Gaps 0;

RESULT 4

S51605

receptor-like tyrosine kinase Ehk-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51605
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:94067777
A;Accession: S51605
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-948 <MAI>
A;Cross-references: EMBL:S68030
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;628-936/Domain: protein kinase homology <KIN>
F;636-644/Region: protein kinase ATP-binding motif

Query Match   68.4%;   Score 1003;   DB 2;   Length 948;
Best Local Similarity   65.7%;   Pred. No. 4.1e-81;
Matches 180;   Conservative 39;   Mismatches 37;   Indels 18;   Gaps 3;

Qy   3 COLSILILLSCSVLDSFGELIP----------QPSNEVNLLDSKTIQGELGWISYPSHGW 52
            ||:| ||||||||||||||||:|||||||||||
Db   4 CEVREFL------QFGFELPLLFAWTGDCSHVSNQVVLLDTSIVMGELGWKTVPLNGW 56

Qy   53 EEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLV 112
            ||:||||||:|| |:||| ||||||||||||||||||:|||||||||||||||||||||:
Db   57 DAITEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIYVEMKFTLRDCNSIPWV 116

Qy   113 LGTCKETFNLIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVG 172
            |||||||||||||||:|| || |:||||||||||||||||||||||||||||||||||||
Db   117 LGTCKETFTLYYIESDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREVG 176

Qy   173 PVNKKGFYLAFQDVGACVALVSVRVRYFKKCPPTVKNLAMFPDTVP-MDSQSLVEVRGSCV 231
            |:|||||||||||||:|||||||||||||||||||||:|||||||| |||||||||||||:
Db   177 PIERKGFYLAFQDIGACIALVSVRVEVKKCPPTVRNIAMFPDTIPRVDSSSLVEVRGSCV 236

Qy   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 265
            |||:||| ||||||||||||||||||||||:|||
Db   237 KSSEERDTPKLLYCGADGDWLVPLGRCICTIGYEE 270

RESULT 5

I78844

receptor protein-tyrosine kinase - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999

C;Accession: I78844
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welcher,
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein-ty
A;Reference number: I58351; MUID:95206782
A;Accession: I78844
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-986 <RES>
A;Cross-references: GB:L36645; NID:g551613; PIDN:AAA74246.1; PID:g551614
A;Genetics:
A;Gene: HEK8
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
F;619-885/Domain: protein kinase homology <KIN>

Query Match
Best Local Similarity 67.9%; Score 995; DB 2; Length 986;
Matches 183; Conservative 37; Mismatches 42; Indels 4; Gaps 3;

```
Qy   10 LLSC--SVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIR 66
Db    9 LFSCLFGICDAVTGSRVYPANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIR 68

Qy   67 TYQVCNVMDHSQNNWLRTNNVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYME 126
Db   69 TYQVCNVMEPSQNNWLRTDWITREGAQRVIEIKFTLRDCNSLPGVWGTCKETFNLYYYE 128

Qy  127 SDDDHGVKFREHQFTKIDTIAADESSTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDV 186
Db  129 SDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDV 188

Qy  187 GACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCS 245
Db  189 GACIALVSVRVFYKRCPLTVRNLAQFPDTITGADISSLVEVRGSCVNNSEEKDVPKMYCG 248

Qy  246 TEGEWLVPIGKCCSCNAGYEERGFMCQ 271
Db  249 ADGEWLVPIGNCLCNAGHEERSGECQ 274
```

RESULT 6
I48967
brain-specific kinase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C;Accession: I48967
R;Zhou, R.R.; Copeland, T.D.; Kromer, L.F.; Schulz, N.T.
J;J. Neurosci. Res. 37, 129-143, 1994
A;Title: Isolation and characterization of Bsk, a growth factor receptor-like tyrosine k
A;Reference number: I48967; MUID:94194581
A;Accession: I48967
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-877 <RES>
A;Cross-references: EMBL:U07357; NID:g463369; PIDN:AAA17038.1; PID:g466370
A;Genetics:
A;Gene: Bsk
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; transmembrane protein
F;512-778/Domain: protein kinase homology <KIN>
F;520-528/Region: protein kinase ATP-binding motif
F;801-868/Domain: SAM homology <SAM>

Query Match
Best Local Similarity 67.5%; Score 990; DB 2; Length 877;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSRTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
Db   46 LLLCAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPPKNGWEEIGEVDENYAPIHTYQ 102
```

RESULT 7
S47489
receptor tyrosine kinase - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C;Accession: S47489
R;Taylor, V.; Pfarr, S.; Miescher, G.C.; Honegger, P.; Breitschopf, H.; Lassmann, H.;
submitted to the EMBL Data Library, April 1994
A;Description: Expression and developmental regulation of EHK-1, a neuronal ELK-like
A;Reference number: S47489
A;Accession: S47489
A;Molecule type: mRNA
A;Status: preliminary
A;Residues: 1-898 <TAY>
A;Cross-references: EMBL:X78689; NID:g531543; PIDN:CAA53357.1; PID:g531544
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;568-834/Domain: protein kinase homology <KIN>
F;576-584/Region: protein kinase ATP-binding motif

```
Qy   70 VCNVMDHSQNNWLRTNNVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
Db  103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy  130 DHGVKFREHQFTKIDTIAADESSTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
Db  163 ENGRSIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy  190 VALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
Db  223 IALVSVRVYIKKCPSVVRHLAIFPDTITGADSSQLLEVSGSCVNHSVTDDPKMHCSAEG 282

Qy  249 EWLVPIGKCCSCNAGYEERGFMCQ 271
Db  283 EWLVPIGKCMCKAGYEEKNGTCQ 305
```

RESULT 8
S51603
receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S51603
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:94067777
A;Reference number:

Query Match
Best Local Similarity 67.4%; Score 988; DB 2; Length 898;
Matches 174; Conservative 44; Mismatches 41; Indels 4; Gaps 2;

```
Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
Db   51 LLLCAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPPKNGWEEIGEVDENYAPIHTYQ 107

Qy   70 VCNVMDHSQNNWLRTNNVPRNSAQKIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
Db  108 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 167

Qy  130 DHGVKFREHQFTKIDTIAADESSTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
Db  168 ENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 227

Qy  190 VALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
Db  228 IALVSVRVYIKKCPSVVRHLAVFPDTITGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 287

Qy  249 EWLVPIGKCCSCNAGYEERGFMCQ 271
Db  288 EWLVPIGKCMCKAGYEEKNGTCQ 310
```

```
A;Accession: S51603
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-893 <MAI>
A;Cross-references: EMBL:S68028
A;Note: the authors translated the codon GAC for residue 170 as Glu
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; transmembrane protein
F;563-829/Domain: protein kinase homology <KIN>
F;571-579/Region: protein kinase ATP-binding motif

Query Match             67.2%;   Score 985;  DB 2;  Length 893;
Best Local Similarity   65.8%;   Pred. No. 1.5e-79;
Matches  173;  Conservative  45;  Mismatches  41;  Indels   4;  Gaps   2;

Qy    10 LLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ 69
         |||:|| | :|||||||: || ||||| ||:| ||||:||||||||||||:||||:|||
Db    46 LLLCAALRT---LLASPSNEVNLLDSRIVLGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy    70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         || ||||||||:|:|||||||||||||:|||||||||||||||||||| |||:|||||||
Db   103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy   190 VALVSVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         || : |||:||||||||||||||||| |||||||||||||||||:||||| ||| |||
Db   223 IALVSVRYIYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
         |||||||| ||||||||||||||  ||
Db   283 EWLVPIGKCMCKAGYEEKNGTCQ 305

RESULT  9
S51604
receptor-like tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C;Accession: S51604
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase fam
A;Reference number: S49015; MUID:9407777
A;Accession: S51604
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-981 <MAI>
A;Cross-references: EMBL:S68029
A;Note: the authors translated the codon GAC for residue 170 as Glu
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
A;Keywords: ATP; transmembrane protein
F;651-917/Domain: protein kinase homology <KIN>
F;659-667/Region: protein kinase ATP-binding motif

Query Match             67.2%;   Score 985;  DB 2;  Length 981;
Best Local Similarity   65.8%;   Pred. No. 1.7e-79;
Matches  173;  Conservative  45;  Mismatches  41;  Indels   4;  Gaps   2;

Qy    10 LLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ 69
         |||:|| | :|||||||: || ||||| ||:| ||||:||||||||||||:||||:|||
Db    46 LLLCAALRT---LLASPSNEVNLLDSRIVLGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy    70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         || ||||||||:|:|||||||||||||:|||||||||||||||||||| |||:|||||||
Db   103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy   190 VALVSVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         || : |||:||||||||||||||||| |||||||||||||||||:||||| ||| |||
Db   163 ENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy   130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         :|: |||||||||||||||||||| :||||||:|||||:||||: :|||||||||||||
Db   163 ENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
         |||||||| ||||||||||||||  ||
Db   283 EWLVPIGKCMCKAGYEEKNGTCQ 305

RESULT  10
S49015
receptor tyrosine kinase Ehk-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S49015; S51602
R;Maisonpierre, P.C.; Barrezueta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A;Title: Ehk-1 and Ehk-2: two novel members of the Eph receptor-like tyrosine kinase
A;Reference number: S49015; MUID:9407777
A;Accession: S49015
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-305, 'G', 359-1005 <MA>
A;Cross-references: EMBL:S68026
A;Note: the authors translated the codon GAC for residue 170 as Glu
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
A;Keywords: ATP; transmembrane protein
F;675-941/Domain: protein kinase homology <KIN>
F;683-691/Region: protein kinase ATP-binding motif

Query Match             67.2%;   Score 985;  DB 2;  Length 1005;
Best Local Similarity   65.8%;   Pred. No. 1.7e-79;
Matches  173;  Conservative  45;  Mismatches  41;  Indels   4;  Gaps   2;

Qy    10 LLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ 69
         |||:|| | :|||||||: || ||||| ||:| ||||:||||||||||||:||||:|||
Db    46 LLLCAALRT---LLASPSNEVNLLDSRIVLGDLGWIAFPKNGWEEIGEVDENYAPIHTYQ 102

Qy    70 VCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
         || ||||||||:|:|||||||||||||:|||||||||||||||||||| |||:|||||||
Db   103 VCKVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD 162

Qy   130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
         :|: |||||||||||||||||||| :||||||:|||||:||||: :|||||||||||||
Db   163 ENGRNIKDNQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 222

Qy   190 VALVSVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
         || : |||:||||||||||||||||| |||||||||||||||||:||||| ||| |||
Db   223 IALVSVRYIYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSCVNHSVTDDPPKMHCSAEG 282

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
         |||||||| ||||||||||||||  ||
Db   283 EWLVPIGKCMCKAGYEEKNGTCQ 305

RESULT  11
I78843
receptor protein-tyrosine kinase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 18-Jun-1999
C;Accession: I78843
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welch
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein
```

A;Reference number: I58351; MUID:95206782
A;Accession: I78043
A;Status type: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-991 <RES>
A;Cross-references: GB:L36644; NID:g551611; PIDN:AAA74245.1; PID:g551612
C;Genetics:
A;Gene: HEK7
C;Superfamily: protein-tyrosine kinase, receptor type eph: fibronectin type III repeat h
F;627-893/Domain: protein kinase homology <KIN>
F;916-982/Domain: SAM homology <SAM>

Query Match   67.1%; Score 984; DB 2; Length 991;
Best Local Similarity   65.8%; Pred. No. 2.1e-79;
Matches 173; Conservative 44; Mismatches 42; Indels 4; Gaps 2;

Qy   10 LLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRTYQ 69
              |||:|||||||||||||||||||||||||||:||||||:|::|||||||||||||||
Db   20 LLICAALRT---LLASPSNEVNLLDSRTVMGDLGWIAFPKNGWEELGEVDENYAPIHTYQ 76

Qy   70 VCNVMDHSQNNWLRTNWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD 129
              |:: |||||| ||:|::||||:|||||||||||||||||||||| |||||||||||||
Db   77 VCKVMEQNQNNWLLISWISNEGASRIFIELKFTLRDCNSLPGVMGTCKETFNMYYFFSDD 136

Qy   130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC 189
              :|:||||||||||||||||:||:||||||||||||||||||||||::|||||||||||
Db   137 QNGRNIKENQIKIDTIAADSFTELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGAC 196

Qy   190 VALVSVRVYEKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG 248
              :|||||:||||:||||||||||||||| |||||||||||||||:||||||||||||||
Db   197 IALVSVRVYYKKCPSVVRHLAVFPDTITGADSSQLLEVSGSCVNHSVTDEPPKMHCSAEG 256

Qy   249 EWLVPIGKCSCNAGYEERGFMCQ 271
              |||||||| |||||||||| ||
Db   257 EWLVPIGKCMCKAGYEEKNGTCQ 279

RESULT 12
S78059
protein-tyrosine kinase (EC 2.7.1.112) Eph precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999
C;Accession: S78059; S30505; I58366
R;Charnay, P.
R:Giardi-Hebenstreit, P.; Nieto, M.A.; Frain, M.; Mattei, M.G.; Chestier, A.; Wilkinson
Oncogene 7, 2499-2506, 1992
A;Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed in t
A;Reference number: S30496; MUID:93096484
A;Accession: S30505
A;Molecule type: mRNA
A;Residues: 1-31,55-986 <GIL>
A;Cross-references: EMBL:X65138
C;Genetics:
A;Gene: Sek
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
C;Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferase;
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-986/Product: protein-tyrosine kinase Eph #status predicted <MAT>
F;548-569/Domain: transmembrane #status predicted <TMM>
F;619-885/Domain: protein kinase homology <KIN>
F;627-635/Region: protein kinase ATP-binding motif
F;900-974/Domain: SAM homology <SAM>
F;235,340,408,423/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match   66.8%; Score 980; DB 2; Length 986;

A;Reference number: I78043
A;Accession: S78059
A;Molecule type: mRNA
A;Reference number: S78059
A;Accession: S78059
submitted to the EMBL Data Library, March 1992

Best Local Similarity   71.4%; Pred. No. 4.7e-79;
Matches 177; Conservative 35; Mismatches 34; Indels 2; Gaps 2;

Qy   26 PSNEVNLLDSKTIQGELGWISYP-SHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
              ||||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Db   27 PANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRYQVCNVMEASQNNWLRT 86

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPF 204
              ||||||:||:||||||||||||||||||||||||||||||||||| |||||||||||||
Db   147 TIAADESFTQVDIGDRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPL 206

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGY 263
              ||||||||||||| ||||||||||||||:||||||||||||| |||||||||||:||||
Db   207 TVRNLAQFPDTITGADISSLVEVRGSCVNSEEKDVPKMYCGADGEWLVPIGNCLCNAGH 266

Qy   264 EERGFMCQ 271
              || ||||
Db   267 EEQNGECQ 274

RESULT 13
I51549
receptor tyrosine kinase - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I51549
R;Winning, R.S.; Sargent, T.D.
Mech. Dev. 46, 219-229, 1994
A;Title: Pagliaccio, a member of the Eph family of receptor tyrosine kinase genes, ha
A;Reference number: I51549; MUID:95001564
A;Accession: I51549
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-985 <WIN>
A;Cross-references: GB:L26099; NID:g4164402; PIDN:AAA64464.1; PID:g4164403
A;Gene: Pag
C;Genetics:
C;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C;Keywords: ATP; transmembrane protein
F;618-884/Domain: protein kinase homology <KIN>
F;626-634/Region: protein kinase ATP-binding motif

Query Match   66.6%; Score 977; DB 2; Length 985;
Best Local Similarity   70.6%; Pred. No. 8.7e-79;
Matches 175; Conservative 39; Mismatches 32; Indels 2; Gaps 2;

Qy   26 PSNEVNLLDSKIIQGELGWISYP-SHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRT 84
              |||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Db   27 PASEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRYQVCNVMESSQNNWLRT 86

Qy   85 NWVPRNSAQKIIVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKID 144
              |:||||||||||||||||||||||||||:||||||||||||::| |||||||||||||
Db   87 DWIPRSGAQRVYVEIKFTLRDCNSLPGVMGTCKETFNLYYESNDKERFIRETQYVKID 146

Qy   145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRYYFKKCPF 204
              ||||||:||||||||||||||:||||||||||||||||||||||| |||||||:|||||
Db   147 TIAADESFTQVDIGDRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPL 206

Qy   205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGY 263
              ||||||||||||| ||||||||||||||:|||||||||||||| |||||||||||:|||
Db   207 .TVRNLAQPPDTIIGSDFSSLVEVRGSCVNSEEKDVPKMYCGADGEWLVPIGNCLCNAGF 266

Qy   264 EERGFMCQ 271
              || ||||
Db   267 EEHNGGCQ 274

RESULT 14

I50615
receptor-type protein-tyrosine kinase Cek7, long splice form - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C;Accession: I50615; I50616; I50614
R;Sleyer, D.A.; Verderame, M.F.
Gene 148, 219-226, 1994
A;Title: identification of a complete Cek7 receptor protein tyrosine kinase coding seque
A;Reference number: I50615; MUID:95047429
A;Accession: I50615
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1013 <STE>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60612.1; PID:g555618
A;Accession: I50616
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-572, 'R', 596-1013 <SI2>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60613.1; PID:g555619
A;Accession: I50614
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-279,444-572,'R',596-1013 <SI3>
A;Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60614.1; PID:g555620
A;Genetics:
A;Gene: Cek7
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat h
A;Keywords: alternative splicing; ATP; transmembrane protein
F;657-915/Domain: protein kinase homology <KIN>
F;649-915/Domain: protein kinase ATP-binding motif
F;938-1004/Domain: SAM homology <SAM>

Query Match　　　64.6%;　　Score 947; DB 2; Length 1013;
Best Local Similarity　　64.6%;　　Pred. No. 4.2e-76;
Matches　170; Conservative　37; Mismatches　52; Indels　4; Gaps　2;

Qy　　10 LISCSVLDSFGELIPQPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRIYQ　69
　　　　||: | ||||| |: |-|-||| ||||:|:||||||||||||||| |||: :|||||||||
Db　　20 LLLCAALRS---LLASPGSEVNLLDSRTVMGDLGWIRAYPKNGWEEIGEVDENYAPIHTYQ　76

Qy　　70 VCNVMDHSQNMWLRTNMVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDD　129
　　　　||:| |||||||||||:|-|||||||||||||||||||||||||||||||:||||||
Db　　77 VCKVMEQNQNMWLLTSWISNEGRPASSFELKFTLRDCNSLPGGLGTCKETFNMYYFESDD　136

Qy　130 DHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGAC　189
　　　　||:|:||||||||||||||||||||||||||:|:|||||||||||||||||||||||
Db　197 VCKVMEQNQNMWLLTSWISNEGRPASSFELKFTLRDCNSLPGGLGTCKETFNMYYFESDD　136

Qy　190 VALVSVRYVFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKEEDPPRMYCSTEG　248
　　　　||||||||||||||:|||||||||||| :||||||||||||:|| ||||||||||||||
Db　137 EDGRNIRENQYIKIDTIAADESFTELDLGDRVWKLNTEVRDVGPLTKKGFYLAFQDVGAC　196

Qy　249 EWLVPIGKCKCSCNAGYERGFMCQ　271
　　　　|||||||| ||||||||||||||
Db　257 EWLVPIGKCLCKAGYEEKNNTCQ　279

RESULT 15
I58351
receptor protein-tyrosine kinase - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
R;Fox, G.M.; Holst, P.L.; Chute, H.T.; Lindberg, R.A.; Janssen, A.M.; Basu, R.; Welcher,
Oncogene 10, 897-905, 1995
A;Title: cDNA cloning and tissue distribution of five human EPH-like receptor protein-ty
A;Accession: I58351; MUID:95206782
A;Reference number: I58351
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-998 <RES>
A;Cross-references: GB:L36642; NID:g551607; PIDN:AAA74243.1; PID:g551608

C;Genetics:
A;Gene: HEK11
A;Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
A;Keywords: ATP; transmembrane protein
F;631-897/Domain: protein kinase homology <KIN>
F;639-647/Domain: protein kinase ATP-binding motif
F;920-986/Domain: SAM homology <SAM>

Query Match　　　63.3%;　　Score 927.5; DB 2; Length 998;
Best Local Similarity　　68.4%;　　Pred. No. 2.2e-74;
Matches　171; Conservative　39; Mismatches　37; Indels　3; Gaps　2;

Qy　　25 QPSNEVNLLDSKTIQGELGWISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNMWLRT　84
　　　　|: |||||||| ||| |-|||| :|||||||||||||||||||||||: :||||||||
Db　　28 QAAKEVLILDSKAQQTELEWISSPPNGWEEISGLDENTPIRIYQVCQVMEPNQNNWLRT　87

Qy　　85 NMVPRNSAQKIYVELKFTLRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKID　144
　　　　|| : :||||||||||||||||||||:|:|| |||||:|||||||:|:|||||||||||
Db　　88 NWISKGNAQRIFVELKFTLRDCNSLPGVLGTCKETFNLYYYEYDYDTGRNIRENLYVKID　147

Qy　145 TIAADESFTQMDLGDRILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRYVFKKCPF　204
　　　　|||||||||:||||||||:|:|:|||||||||||||||||||||||||||||||||||
Db　148 TIAADESFTQGDLGERKMKLNTEVREIGPLSKKGFYLAFQDVGACIALVSVKYYKKCWS　207

Qy　205 TVKNLAMFPDTVP-MDSQSLVEVRGSCVNNSKE--EDPPRMYCSTEGEWLVPIGKCKCSCNA　261
　　　　||||||||||| : :|||||||||||:||:||: | |: ||||||||||||||| |
Db　208 IIENLAIFPDTVTGSEFSSLVEVRGTCVSSAEEEAENAPRMHCSAEGEWLVPIGKCICKA　267

Qy　262 GYEERGFMCQ　271
　　　　|:||:|||:
Db　268 GYQOKGDTCE　277

Search completed: May　9, 2000, 22:27:20
Job time: 4159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 06:01:17 ; Search time 1358.94 Seconds
                 (without alignments)
                 2424.883 Million cell updates/sec

Title: US-09-104-340-5
Perfect score: 813
Sequence: 1 atggattgtcagctctccat.........aaagaggtttatgtgccaa 813

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est1:*
21: em_est1:*
22: em_est3:*
23: em_est4:*
24: em_est4:*
25: em_est5:*
26: em_est6:*
27: em_est7:*
28: gb_est8:*
29: gb_est9:*
30: gb_est10:*
31: gb_est11:*
32: gb_est12:*
33: gb_est13:*
34: gb_est14:*
35: gb_est15:*
36: gb_est16:*
37: gb_est17:*
38: gb_est18:*
39: gb_est19:*
40: gb_est20:*
41: gb_est21:*
42: gb_est22:*
43: gb_est23:*
44: gb_est24:*
45: gb_est25:*

46: gb_est26:*
47: gb_est27:*
48: gb_est28:*
49: gb_est29:*
50: gb_est30:*
51: gb_est31:*
52: gb_est32:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: gb_gss1:*
87: gb_gss2:*
88: gb_gss3:*
89: gb_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: gb_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

%
Result   Query

No.	Score	Match Length	DB	ID	Description
1	349	42.9	435	23	H38363	H38363 yp50g04.r1
2	297.6	36.6	835	69	AU080921	AU080921
3	201.6	24.8	572	74	AW175051	AW175051 f131g02.y
4	147.4	18.1	474	41	AI011631	AI011631 EST206082
5	145	17.8	473	41	AI011632	AI011632 EST206083
6	143.4	17.6	424	84	B36983	B36983 HS-1042-A2-
7	126.0	15.6	406	44	AI325344	AI325344 m101c02.y
8	119.2	14.7	402	20	R15219	R15219 yf88e12.r1
9	112.2	13.8	493	44	AI249967	AI249967 qx47d02.x
10	110.2	13.6	366	27	AA024123	AA024123 m101c02.r
11	103.2	12.7	348	45	AI323334	AI323334 m101c02.x
12	92.6	11.4	496	45	AI337308	AI337308 tb98b01.x
13	89	10.9	314	64	AW045268	AW045268 UI-M-BH1-
14	88.2	10.8	418	61	AI828089	AI828089 wk3if07.x
15	86	10.6	344	80	AW366937	AW366937 ILO-HT015
16	83	10.2	572	62	AI878971	AI878971 au51h08.y
17	80.2	9.9	221	23	AL079781	AL079781 DKFZp434M
18	75.4	9.3	296	23	H25851	H25851 yl53g03.s1
19	70.2	8.6	400	79	AW233440	AW233440 fj32f03.y
20	68.6	8.4	628	43	AI003549	AI003549 ai88a09.s
21	68.6	8.4	800	61	AL040428	AL040428 DKFZp434C
22	64	7.9	474	61	AL045967	AL045967 DKFZp434N
23	62.8	7.7	687	61	AI828089	AI828089 wk3if07.x
24	62.2	7.7	375	21	T78739	T78739 yd01e03.r1
25	60.4	7.4	252	25	N87424	N87424 L386SF Huma
26	54	6.7	151	42	AI794583	AI794583 fc4b08.y
27	50.8	6.2	416	103	AF071057	AF071057
28	50.2	6.2	643	103	AA201571	AA201571 RCCII1-58
29	40.2	4.9	1099	83	AL104186	AL104186 Drosophil
30	39.8	4.9	494	59	AI767989	AI767989 wj01dl1.x
31	39.8	4.9	1063	83	CNS01517	CNS01517 Drosophil
32	39	4.8	552	60	AL119882	AL119882 Drosophil
33	37.4	4.6	324	20	M79194	M79194 EST01342 Su
34	36.2	4.5	802	105	AQ330023	AQ330023 nbxb0046K
35	35.6	4.4	566	108	AQ526929	AQ526929 CITB1-E1-
36	35.4	4.4	237	84	B47312	B47312 HS-1056-A1-
37	35.2	4.4	1101	38	CNS0182P	CNS0182P Drosophil
38	35	4.3	333	63	AW017395	AW017395 nx69d02.s
39	34.8	4.3	462	63	AW746914	AW746914
40	34.8	4.3	637	79	AW556663	AW556663 EST304800
41	34.6	4.3	459	45	AI366919	AI366919 qv94b06.x
42	34.4	4.2	629	42	AI143874	AI143874 qe04f05.x
43	34	4.2	320	42	AI112217	AI112217 UI-R-Y0-m
44	34	4.2	1101	82	CNS0106X	CNS0106X Drosophil
45	33.8	4.2	246	94	AQ036415	AQ036415 CIT-HSP-2

RESULT 1

```
LOCUS       H38363                   435 bp    mRNA            EST     16-AUG-1995
DEFINITION  H38363 yp50g04.r1 Soares retina N2b4HR Homo sapiens cDNA clone
            IMAGE:190902 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
            RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   H38363
VERSION     H38363.1  GI:907862
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 435)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On May 5, 1995 this sequence version replaced gi:797807.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2663
            High quality sequence stops: 276
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2663    Std Error: 0.00
            Seq primer: M13Rev
            High quality sequence stop: 276.
FEATURES             Location/Qualifiers
     source          1..435
                     /organism="Homo sapiens"
                     /db_xref="GDB:3847311"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:190902"
                     /clone_lib="Soares retina N2b4HR"
                     /sex="male"
                     /tissue_type="retina"
                     /dev_stage="55 year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCGCGCCGCGGCCGCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). The retinas were obtained from a 55 year old
                     Caucasian and total cellular poly(A)+ RNA was extracted 6
                     hrs after their removal. The retina RNA was kindly
                     provided by Roderick R. McInnes M.D. Ph.D. from the
                     University of Toronto. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      109 a    108 c    103 g    109 t      6 others
ORIGIN
```

ALIGNMENTS

```
Query Match     42.9%;  Score 349;  DB 23;  Length 435;
Best Local Similarity   97.5%;   Pred. No. 6.4e-96;
Matches 352;  Conservative   0;  Mismatches   9;  Indels   0;  Gaps   0;

Qy    3 ggattgtcagtctccatcctcctcctcttctcagctgtctgtctgtctctgctcagcttcggggga 62
          |||  ||||||||||||| ||| | ||| ||||||||| ||  ||||||| ||| ||||||||||
Db   69 GCAATGTGCAGCTCTCCATCCTCCCTCTCCCTTCCAGCTGCCTCCGTTTCTGGACAGCTTCGGGGA 128

Qy   63 actgattccgcagccttccaatgaagtcaatctactggattcaaagaaaacaaattcaaggga 122
         || ||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||||
Db  129 ACTNAATTCCGCAGCCTTCCANATGAAGTCAATCTACTGGATTCAAAAAACAAATTCAAGGGA 188

Qy  123 gctgggctggatctcttatccatcacacatgggtggaagagatacatgtggtggatgaaca 182
         ||||||||||||||||||||||||||||||||||| |||||||||| |||||||||||||||
Db  189 GCTGGGCTGGATCTCTTATCCATCACACATGGGTGGGAAGAGATACAGTGTGGTGGATGAACA 248

Qy  183 ttacacacaccatcaggacttaccaggtgtgcaatgtcatggaccaacagtcaaaacaattg 242
         ||||| | |||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  249 TTACACACCATCAGGACTTACCAGGTGTGCAATGTCATGGACCAACAGTCATAACAATTG 308

Qy  243 gctgagaacaaactgggtccccaggaactcagtcagagagattatgtggagctcaagtt 302
         ||||||||||| ||||||||||||||||||| |||||||||| |||||||||||||||||
Db  309 GCTGAGAACAAAACTGGGTCCCCAGGAACTCAGCTCAGAGAGATTTATGTGGAGCTCAAGTT 368

Qy  303 cacttctacgagactgcaatagcattcctttggttttaggaacttgcaaggagagacattcaa 362
         ||||| ||||||||||||||| |||||| || ||| |||| |||| |||| |||||||| ||
Db  369 CACTCTACGAGACTGCAATAGCAATTCCANTTNGTTTAGGNACTTNCAAGGGACATTAAC 428

Qy  363 c 363
          |
Db  363 c 363
```

Db   429 c 429

RESULT 2
AU080921
LOCUS   AU080921   835 bp   mRNA   EST   20-OCT-1999
DEFINITION AU080921 Sugano mouse brain mncb Mus musculus cDNA clone MNCb-6241
            5', mRNA sequence.
ACCESSION   AU080921
VERSION   AU080921.1 GI:6085675
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 835)
AUTHORS   Hashimoto,K., Kusuda,J., Toyoda,A., Tanuma,R., Ito,A., Hirata,M.,
            Suzuki,Y., Sasaki,M. and Sugano,S.
TITLE   Isolation of full-length cDNA clones from a mouse brain cDNA
            library made by oligo-capping method
JOURNAL   Unpublished (1999)
COMMENT   On May 18, 1998 this sequence version replaced gi:3137227.
            Contact: Katsuyuki Hashimoto
            National Institute of Infectious Diseases
            Division of Genetic Resources
            23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            Email: khashm@nih.go.jp
            URL: http://www.nih.go.jp/yoken/genbank/.
FEATURES   Location/Qualifiers
source   1..835
            /organism="Mus musculus"
            /strain="C57BL"
            /db_xref="taxon:10090"
            /clone="MNCb-6241"
            /clone_lib="Sugano mouse brain mncb"
            /sex="female"
            /dev_stage="adult"
            /lab_host="top10"
            /note="Organ: brain; Vector: pME18S-FL3; 1st strand cDNA
            was primed with an oligo(dT) primer
            [ATGTGGCCTATTTTTTTTTTTTTTTTTT] double-stranded cDNA was
            ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
            and cloned into distinct DraIII sites of the pME18S-FL3.
            XhoI sites just outside the DraIII sites can be used to
            isolate the cDNA insert. Size selection was performed to
            exclude fragments <1.5 kb. Library was constructed by
            Sugano et al. (University of Tokyo, Institute of Medical
            Science). Custom primer for sequencing: 5' end primer
            [CTTCTGCTCTAAAAAGCTGCG]"
BASE COUNT   202 a   230 g   190 t   31 others
ORIGIN

Query Match   36.68;   Score 297.6;   DB 69;   Length 835;
Best Local Similarity   63.2%;   Pred. No. 4.2e-80;
Matches 469;   Conservative 0;   Mismatches 265;   Indels 8;   Gaps 2;

Qy   16 tccatcctcctccctcgctgctctcgctcggggaactgattccgcag 75
          ||||||| | || | || ||| | ||| || | | |||| |
Db   75 TTCATCCTCTTCTTCGTTCTCTTTGGAATTTGCGACGCTGTCACCGGTTCTCAGGGTATAC 134

Qy   76 ccctccaatgaagtcaatctactgattccaaaacaattcaaggggagctgggctggatc 135
          |||| |||||||||| | | ||| ||||||||| || |||||||||||| ||| |||||
Db   135 CCCCGGAATGAAGTTACTTTATTTGGATTCCAGARTCTGTTCACGGAGAGCGTTGGGGTGGGATA 194

Qy   136 tcttatccat--cacatgggtgggagagatcagtggttgtggtgtgaacacattacacacc 192
          | | |||| |||||||||||| ||||| |||||| ||||| |||| ||| ||||| |||
Db   195 GCAAGCCCTCTGGAAGGAGGGGTGGGGGAGGAAGTGACCCATTATGATGGAGAAAAATACACCG 254

Qy   193 atcaggaacttaccaggttgtgcaaatgtcaatgtcatgggaccaacagtcaaaaacaaatggctgaagaaca 252
          ||| | |||||||||||| |||||| || ||| | | || | || |||| ||||| | |||| ||| |
Db   255 ATCCGAACCTACCAGGTGTGCAACGTGATGGAAGCCAGCCAGACAACAACTGCCTGCGAACT 314

Qy   253 aactggtgtcccaagaactcaagctcaagaaagatttatgtgtggaagctcaactctacga 312
          |||||| | |||| ||| || ||| |||||||||||| ||||| |||||||||||||||
Db   315 GACTGGGATCACCCGCGAAGGGGCGCGCAGAGGGTGTACATTGAAATTAAGTTCACTCTAAGG 374

Qy   313 gactgcaatagcattccattggtttaggaacttgcaaggagaagacattcaaacctgtaactac 372
          |||||||| || ||||| ||| | || | | ||||||| ||| ||| |||| |||||||||
Db   375 GACTGCAACAGTCTTCCGGGCGTCATGGGCGGTCATGGGAGGACTTCCAAGGAGACGTTAACCTCTACTAC 434

Qy   373 atgggagtcctgatcatgggtgaaattcgaagttacagttacaaagattgac 432
          || |||||| |||||||||||| |||||| ||||||| ||||||||||||||
Db   435 TATGAGTCGGACAACGACAAGGACGCCCTTCATCAGAGACAGATCGAC 494

Qy   433 accattgcagctgatgaaagttcactcaatgatcttgggaccgtattctgaagctc 492
          |||| | ||||| ||||| ||| ||| | | | ||||||||| ||| | |||||
Db   495 ACCATGCCAGCTGCGATGAGAGCTTCACACGCGGTGGACATTGGTGACCGAAATCATGAAAACTC 554

Qy   493 aacactgagattagagaagtaggtcctgtcaacagagaaggggatttcattgtcattcaa 552
          ||||||||||| | | | ||| | ||| |||||| ||||| || ||| || | ||||
Db   555 AACACTGAGATCCGGGATGTGAGGGCCACTGAGCAGACAAGAAGGGGTTTTAACCTGGCTTTCAG 614

Qy   553 gatgttggtgcttgttgtccttggtgtcttgtgagagtatacttcaaaaagtgcccattt 612
          ||||| || ||||| | ||| | ||||||| | ||| ||||| |||||| |||| |||
Db   615 GATGCTGGGTGCTTGCATCCGCCTCGGTGTCTCCNTGTTCTACAANAAGTGNCCACTT 674

Qy   613 acagtgaagaatctggctatgtttccaacaacggtaccccatggact----cccagtccc 667
          || |||| |||| ||| || ||||||| | ||| | | || | ||| |||| |||
Db   675 NCCCGGTTCNAAATTANCCCANTTTCCTGATACCATTACTGGGGCTGGATGNCCNAAAATGN 734

Qy   668 tggtggaggttagagggtctctgtgtcaacaattctaaggaggagatcctccaaggatgt 727
          ||| |||| || ||| |||| ||| | |||||| | | | || ||| | ||| | |
Db   735 TGGNGGAAGTTCNAGGCCTCCTTGNGTNAACNACTNANNAANAAGGATGNGCCNAAAATGN 794

Qy   728 actgcagtacaagagcggaatg 749
          || | || | ||| ||| ||
Db   795 NCTGNGGGGCNNANGNNANGNNANG 816

RESULT 3
AW175051
LOCUS   AW175051   572 bp   mRNA   EST   16-NOV-1999
DEFINITION f131g02.y1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
            2639282 5' similar to SW:EPA4_CHICK Q07496 EPHRIN TYPE-A RECEPTOR 4
            PRECURSOR ;, mRNA sequence.
ACCESSION   AW175051
VERSION   AW175051.1 GI:6441079
KEYWORDS   EST.
SOURCE   zebrafish.
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
            Cyprinoidea; Cyprinidae; Rasborinae; Danio.
REFERENCE   1 (bases 1 to 572)
AUTHORS   Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hiller,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B.,
            Ritter,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE   Washu Zebrafish EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3189572.
            Other_ESTs: f131g02.x1
            Contact: S. L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami
            Sequencing by: Washington University Genome Sequencing Center
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 485.
FEATURES   Location/Qualifiers
source   1..572
            /organism="Danio rerio"

RESULT 4

LOCUS   AI011631   474 bp   mRNA   EST   31-JAN-1999
DEFINITION   EST206082 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
             ROVAX55 3' end, mRNA sequence.
ACCESSION   AI011631
VERSION   AI011631.1   GI:3225463
KEYWORDS   EST.
SOURCE   Rattus sp.
  ORGANISM   Rattus sp.
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 474)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
             Kerlavage,A.R. and Adams,M.D.
  TITLE   Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
             Gene Index
  JOURNAL   Unpublished (1998)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2153373.
             Contact: Lee, NH
             ATCC
             The Institute for Genomic Research
             9712, Medical Center Drive, Rockville, MD 20850, USA
             Tel: (301)-838-3529
             Fax: (301)-838-0208
             Email: nhlee@tigr.org
             Seq primer: M13-21.
FEATURES   Location/Qualifiers
     source   1..474
                     /organism="Rattus sp."
                     /db_xref="ATCC (inhost):2021827"
                     /clone="ROVAX55"
                     /db_xref="taxon:10118"
                     /note="Normalized rat ovary, Bento Soares"
                     /clone_lib="Normalized rat ovary, Bento Soares"
                     /organism: ovary; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI"
BASE COUNT   117 a   129 c   135 g   93 t
ORIGIN

Query Match   18.1%;   Score 147.4;   DB 41;   Length 474;
Best Local Similarity   57.1%;   Pred. No. 2.8e-34;
Matches 268; Conservative   0; Mismatches 201; Indels   0; Gaps   0;

Qy   71 cgcagccttccaatgaagtcaatcctactggattcaaaaaacaattcaagggaggagctggggct 130
          ||||||||||||| |||||| | |||||||||| ||||||| || |||||||||||| |||||
Db   4 CGGAACAGGCAAAAGAAGAGTTGTTTTGCTGGACTTCGCAGCAATGAAGGGGAACTCGGCT 63

Qy   131 ggatcttcatccatacaatgggtgggaagagagatcagtggtgtgggatgaacattacac 190
          || || ||||||||| |||| |||| |||||||| ||||||| ||||| || |||||||
Db   64 GGCTCACACACCCCTACGGAAAAGGGTGGGAACCTGATGCAAAACATCATGAATGACATGC 123

Qy   191 ccatcaggacttaccaggtgtgcaatgtcatggaccacagtcaaaacaattggctgagaa 250
          ||||||| |||| || ||| |||||| |||| | ||||||| || ||||| |||| ||
Db   124 CTATCACATGACTCCGTGTGCAAATGTGGTGTCTGCGGACCAGGACAAACTGGCTCGGTA 183

Qy   251 caaactgggtcccagagaagctcagctcagagagatttatgtgtggagctcaagtcactctac 310
          ||| ||||||| ||| ||||||||||||||| |||| ||||| |||||| ||||| |||||
Db   184 CCAACTGGGTGTGTACCGGGAGAGGCCGAGCGCATCGTTATTGAGCTCAAGTCACCGTGC 243

Qy   311 gagactgcaatagccattccaatggtttcaggaggagcatcagttcaacctgtact 370
          ||||||||||| | || ||| ||||| ||| |||||| || || || |||||||
Db   244 GAGACTGCAATAACCTTTCCCGGTGGGCCGCCAGCGTCCTCCAAAGAGACCTCAACCTCTACT 303

Qy   371 acatggagtctgatgatgatcatgggtgaaatttcgagagcatcagttacaaagattg 430
          || |||||| ||||| | |||| ||||| |||| |||||||| ||| || |||||
Db   304 ATGCAGAGTCAGACGTGGACTAGGCACCAACTTCCAGAAGCGCACGTTCACCAAGATCG 363

Qy   431 acaccattgcagctgatgaaagtttcactcaaatggatcttggggaccgtgttctgaagc 490
          || |||||||||| ||||||||||||| ||||| || |||||||||| ||| |||
Db   364 ACACCATCGCCCCTGACGGAGATCACGGTCAGCAGTGACTTCGGAGGCCTGCAATGTCAAGC 423

Qy   491 tcaacactgagattagagaagtaggtcctgtcaacaagaaggagggatttta 539
          || ||| ||||||| ||||||| || ||| ||| ||||||||| |||||
Db   424 TGAATGTGGAGAGGAGCGCATGGTGGGGCGCCCTTACCCGAAGGGCCTTCTA 472

RESULT 5

LOCUS   AI011632   473 bp   mRNA   EST   31-JAN-1999
DEFINITION   EST206083 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
             ROVAX56 3' end, mRNA sequence.
ACCESSION   AI011632
VERSION   AI011632.1   GI:3225464

BASE COUNT   169 a   136 c   160 g   104 t
ORIGIN

Query Match   24.8%;   Score 201.6;   DB 74;   Length 572;
Best Local Similarity   68.6%;   Pred. No. 8e-51;
Matches 306; Conservative   0; Mismatches 136; Indels   4; Gaps   2;

Qy   74 agccttccaatgaagtcaatctactggattcaaaaaacaattcaagggaggagctgggctgga 133
          | ||| |||||||||||||| ||| ||||| |||||| |||||||| || ||||||| |||
Db   113 ACCCTCCTAATCAAGTGACACTGCTCGAACTGTCGCAAGGAGAGAATTGAAATGGG 172

Qy   134 tctcttatccatc--acatgggtgggaagagagatcagtggtggatgaacattacac 190
          | ||| |||||| | |||||||||||||| || ||||| ||| || || |||||||
Db   173 TTGCAAGCCCACGGGAGGGAGGGTGGGAAGAGAGGTGAGCATCATTGGAGGAGGATACGC 232

Qy   191 ccatca-ggacttaccaggtgtgcaatgtcatggaccacagtcaaaacaattggctcactcta 249
          ||||| ||||| || ||| ||| |||||||||||| ||||||| || ||||||||| ||
Db   233 CAATCAGGGAGGACGTCCCANGTGTTGCAAGGTAATAGAACCCAGCACAACTGGCTCTCCGA 292

Qy   250 acaaactgggtccccaggaatctcagctcagaagagattatgtgagctcaagttcactcta 309
          || ||||||||| ||| || ||||||||| |||||| |||||| |||||| ||| ||||
Db   293 ACGGACTGGGATCCCCGTGGCGGAGCTTCAGCCGGGTCTACATCGAGATCAAAAATTCACCCTG 352

Qy   310 cgagactgcaatagcattccatctggtgtttttaggaacttgcaaggagacaattcaacctgtac 369
          || |||||||| | || ||| |||| || |||||||| ||| ||||||| ||||||| |||||
Db   353 CGTGACTGCACAGCCGTGCCAGGAGTGCATGGGAACCTGCGAAGGAGACCTTAATCTTTAC 412

Qy   370 tacatggagtctgatgatgatcaagtgaatacaagtttacaaagatt 429
          || |||||| ||||| | |||||| |||||||||| ||||||||||||
Db   413 TATTATGAGTCCAACAATGACAAAGACGGATACATCCGAGAAAACCAGTTCACCAAAAATC 472

Qy   430 gacaacaattgcaagctaatgaagtttcactcaaatggatcttgggaccgtattctgaag 489
          || |||| ||||||||| || || |||||| || |||||||||||||| | |||||||
Db   473 GACACCANTTGCAGCTGATGAAAGCTTCACGCAGGTGACCATCAGGTCAGCAATCATGAAG 532

Qy   490 ctcaacctgagattgagaagtagg 515
          |||||| || || || ||||| ||
Db   533 CTGAACACNGAGGTGAGGGATGTGGG 558

RESULT 6

```
LOCUS       B36983                424 bp    DNA        GSS       17-OCT-1997
```

KEYWORDS   EST.
SOURCE   Rattus sp.
  ORGANISM Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 473)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
  TITLE   Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
            Gene Index
  JOURNAL   Unpublished (1998)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2153374.
            ATCC
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Contact: Lee, NH
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            Seq primer: M13-21.
FEATURES   Location/Qualifiers
     source   1..473
                     /organism="Rattus sp."
                     /db_xref="ATCC (inhost):2021828"
                     /db_xref="taxon:10118"
                     /clone="ROVAX56"
                     /clone_lib="Normalized rat ovary, Bento Soares"
                     /note="Organ: ovary; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI"
BASE COUNT   116 a   127 c   136 g   93 t   1 others
ORIGIN

Query Match   17.8%; Score 145; DB 41; Length 473;
  Best Local Similarity   57.3%; Pred. No. 1.5e-33;
  Matches 262; Conservative   0; Mismatches 195; Indels   0; Gaps   0;

```
Qy     82 aatggaagtcaatctactggttcaaaaacaacaattcaaggggggagctggtcggatctcttat 141
             ||| | || ||||  |||| || |    |||||||||||||  |||||||| || ||| || |
Db     17 AAAGAAGTTGTTTTGCTGGACTTCGCAGCAATGAAGGGGAACTCGGCTGCCTGGCTCACAC 76

Qy    142 ccatcaacatgggtgggaagagatcagtggtggtggataacattacacaccatccaggact 201
             ||   ||||| ||| |||| |||||||  |  || ||| |||||||||| |||| || |||
Db     77 CCCTACGGAAAAAGGGTGGGACCTGATCGAAAACATCATGAATGACACCTATCTACATG 136

Qy    202 taccaggtgtgcaatctcatggaccacagtcaaaaacaaactgggtc 261
             ||| |||||||||| ||| |||| |||||| |  |||||||||||
Db    137 TACTCCGTGTGCAAATGTGGTGTTCTGGCGACCAGGACAACTGGGTG 196

Qy    262 cccaggaactcagctcagaagagttatgtgggagctcactcacgagactgcaat 321
             |||  ||||||| || || ||||||||| ||||||||| || ||| ||||||||
Db    197 TACCGGCGGAGGAGGCCCAGCGATCTTTATTGAGCTCAAAGTCACCTGCGAGACTGCAAT 256

Qy    322 agcattccattggttttaggaacttgcaaggagacattcaacctgtactacatggagtct 381
             |||  ||||  ||| |||||| |||| |||| ||| | ||  | ||| ||||| |||| |
Db    257 AGCTTTCCGGGGTGGCGGCCCAGCCTCCGAAAGAGAGACCTCTACTATCCAGAGTCA 316

Qy    382 gatgatgatcatgggtgaaatttcgagagagatcatttacaaagattgaaccattgca 441
             ||| ||| || ||| ||||| |||| || |||||| || ||| ||||| ||||||||
Db    317 GACGTGGACATATGGCACCAACTTCAGAAGCGGCCAGTTCACCAAGAATCGACACCATCGCC 376

Qy    442 gctgatgaaagtttcactcaaatggatctctggggaccgtattcctgaagctcaaactgag 501
             ||||| ||||| ||||| |||| |||| ||||| ||||| || || || |||||||| |||
Db    377 CCTGACGAGAATCACGGTCAGCAGTGACTTCAAGGGCTTCAAGGCTCAAACGTGAAATGT 436

Qy    502 attagaagaagtaggtcctgtcaacaagaagggattt 538
             || || |||| | |||||||| ||||||||||| |||
Db    437 GAGCGCATGGTGTGGGCGCCCCTTACCCGGAAGGCTTCT 473
```

RESULT 7

```
LOCUS       AI325344/c            406 bp    mRNA                 23-DEC-1998
DEFINITION  mi01c02.y1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
```

ACCESSION   AI325344
DEFINITION HS-1042-A2-F05-MR.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 824 Col-10 Row=K, genomic survey
            sequence.
VERSION   B36983
ACCESSION   B36983
VERSION   B36983.1   GI:2536352
KEYWORDS   GSS.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 424)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE   Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 824 row: K column: 10
            Class: BAC ends
            High quality sequence stop: 424.
FEATURES   Location/Qualifiers
     source   1..424
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 824 Col-10 Row=K"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /sex="M"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT   77 a   101 c   122 g   122 t   1 others
ORIGIN

Query Match   17.6%; Score 143.4; DB 84; Length 424;
  Best Local Similarity   67.7%; Pred. No. 4.0e-33;
  Matches 233; Conservative   0; Mismatches 102; Indels   9; Gaps   2;

```
Qy    479 gtattctgaagctcaacactgagattaggagaagtaggtcctgtcaacaagaagggatttt 538
            | | ||  |||||||| |||||| || | |||| ||| |||||| |||||||||| ||||
Db      2 GNATACTCAAGCTTAACACTGAGGTGAGGAGAATTGGACCTTTGTCCAAAAAGGGATTCT 61

Qy    539 atttggcatttcaagatgttgttgcctgttgttgcctggtgtctgttgagagtatacttca 598
            || ||||| ||| ||||||||| |||| |||| || ||| |||| || ||||| | |||||
Db     62 ATCTTGCCTTTCAGGATGTTAGGGCGTTCCATACGTTTTGGTTCTTGTCAAAAGTGTACTACA 121

Qy    599 aaaagtgcccatttacagtgaagaaatctggctatgtgttcccagacacacggtacc---catgg 655
            |||||| ||| |  ||||| ||||| |||| || ||||||||||||| || | |||   ||||
Db    122 AGAAGTGCTGGGTCCATTAATTGAGAACTTAGCTATCTTTCCAGATACGACTGGTTGGGTTCAG 181

Qy    656 actcccagtccctggagggttagaggtcttgtcaacaattctaagga-------gg 709
            || |||| || |||||||| |||| ||||| |||||||| || |||       ||
Db    182 AATTTTCCTCTTTAGTCGAGGTTCGAGGTGCAGCAGTGCAATGTGTCAGGGAGGAAGGG 241

Qy    710 aagatctccaaggatgtcactgtctccagtacgaaaggcgaatggcttgtccattggcaagt 769
            ||||||||||| || |||||| |  |||| |  ||||| ||| ||||| ||||| |||| |||
Db    242 AAAACGCCCCAGGAATGCCACTGCAGTGCGAGGAGAGAAGGTTAGTGCCCATTGGGAAAAT 301

Qy    770 gttcctgcaatgctggctatgaagaagagagtgttttatgtgcaa 813
            |  ||||| |||| |||| | ||||||| | |||  |||||||||
Db    302 GTATCTGCAAACGAGGCTACCAGCAACAAAAGAAGGAGACACTTGTGAA 345
```

```
ACCESSION   AI325344
VERSION     AI325344.1  GI:4059773
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2152238.
            Contact: Marra M/Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276058
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1..406
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:459170"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGACGGCGCCGGAAATTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       79 a       117 c       109 g       100 t       1 others
ORIGIN
```

```
Qy      82  aatgaagtcaatctactggattcaaaaacaaattcaaggggagctgggctgggcttat  141
            || |||||  || ||||| |||||||||||||||||||||||||||||||||||||||
Db     393  AAGGAAGTGTTTTGTTGGACTTCCGCACACAATGAAGGGAGCTGCGGGCTCACCGCAC  334

Qy     142  ccatcacatgggtgggggaagagatcaagtggtgtggaacattacacaccatccaggact  201
            ||  | |||  |  |||   | ||||| |||| ||  || |||| | ||||| |||||||
Db     333  CCCTATGGCAAAGGGTGGGCAACCTGATGCAACATCATGGACAGCCTATCTACATG  274

Qy     202  taccaggtgtgcaatgtcatggaccaccagtcaaaacaattggctgagaacaaactgggtc  261
            || |||||||| ||| | |||||  ||||| |  ||||| || |||| | |||||||||
Db     273  TACTCGGTGTGTGCAACGTGGTGTATCCGGCGACCAGGACACAACTGGCTCCGGCCACCAAC  214
```

```
RESULT   8

LOCUS       R15219                402 bp    mRNA              EST        13-APR-1995
DEFINITION  yf88e12.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:29543 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
            RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   R15219
VERSION     R15219.1  GI:769492
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 402)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On Sep 21, 1992 this sequence version replaced gi:276397.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1695
            High quality sequence stops: 337 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1695  Std Error: 0.00
            Seq primer: M13Rpl
            High quality sequence stop: 337.
FEATURES             Location/Qualifiers
     source          1..402
                     /organism="Homo sapiens"
                     /db_xref="GDB:401890"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:29543"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAANTCGGCCGCGGCCGAGGAANTTTTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and Hind III and
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      108 a        75 c       110 g       108 t       1 others
ORIGIN
```

```
ACCESSION   AI325344
   clone IMAGE:459170 5' similar to gb:U07634 Mus musculus
   receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
```

```
Qy     262  cccaggaacttcagcttcagaagatttatgtggagctcaagttccacctctacgagacctgcaat  321
            || ||| |  ||| |||| || |||| ||| | |||| | ||| |||||| || || ||  ||
Db     213  TACCGGGAGGAGGCCGAGCGCATCTTATTATTGAGCTCAAGTTCACGGGTGCGAGACTGTGAAC  154

Qy     322  agcattccattggttttaggagaacttgcaaggagagacatccaacctgtactacatggtgtct  381
            || |||||  |  ||| ||| |||   |   |  |||| ||   ||| || || || |||||||
Db     153  AGCTTCCCGGGGGTGCGCCAGCTCAAGGAGAGACCTTCAAACCTTCACTATGCGAGACTCA  94

Qy     382  gatgatgatcatgggtgaaattccgagagcatcagttacaaagatgaccaccattgca  441
            || ||| ||| |||||  |||| | | |||| ||| |  |||||| |||||||||||
Db      93  GATGTGGACTATGCCACCAACTTCCAGAAGCGCCAGTTCACCAAGATTGACACCACCATCCCC  34

Qy     442  gctgatgaaa  451
            || || ||||
Db      33  CCTGACGAGA  24
```

```
BASE COUNT       79 a       117 c       109 g       100 t       1 others
ORIGIN
```

Query Match   15.6%; Score 126.8; DB 44; Length 406;
Best Local Similarity 58.9%; Pred. No. 5.2e-28;
Matches 218; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

RESULT 9

```
Query Match      14.7%;   Score 119.2;  DB 21;  Length 402;
Best Local Similarity   67.9%;  Pred. No. 1.1e-25;
Matches 199; Conservative  0; Mismatches  84; Indels 10; Gaps  2;

Qy  485 tgaagctcaacactgagaattagagaagtaggtcctgtcaacaagaaggattttattgg 544
        ||||||||||||| ||||||||||||||||||||| |||| ||||||| |||  ||| |
Db    6 TGAAGCTTAAACACTGAGGTGAGGATTGAGACCTTGTCAAGAAGAGGATTCTATCTTG 65

Qy  545 catttcaagatgttggtcttgttgttcctgtctgtgaggtactactcaaaagt 604
        |||||| ||||||| |||| ||||||||| |||||||| || || | |||||
Db   66 CCTTTCAGGATGTTAGGGCCTTGCATAGCTTTGGTTTCGTCAAAGTGTTACTACAGAGT 125

Qy  605 gcccattacagtgaagaatctggctatgttcccagacaacgtacc---catggactccc 661
        |||  |  | | || |  ||| |  ||| |||  | |  | ||    |||  | | ||
Db  126 GCTGGTCCATTATTGAGAACTTAGCTACTCTTTCCAGATACAGTGACTGGTTCAGAATTT 185

Qy  662 agtccctggtgaggttagaagggtcttgtgtcaacaattctaggaggaag------at 714
        |||  |  |  ||| | || ||| | |||||| |  ||  | ||||| |      ||
Db  186 CCTCTTTAGTCGAGGTTCGAGGGACACATGTGTCAGCAGTGCAGGAGAAGCGGAAAACG 245

Qy  715 cctccaaggagttgactgcagtacaagagcgaatggcttgttacccattggcaa 767
        || ||| |||| ||| ||||  ||||| ||  ||| | || |||||||||||
Db  246 CCCCCAGGGATGCGACTGCGTCAGNAAGGAGAAGTGGTTAGTGCCCATTGGGAA 298
```

LOCUS   AI249967   493 bp   mRNA   EST   05-NOV-1998
DEFINITION qx47d02.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2004483 3'
            similar to gb:M59371 TYROSINE-PROTEIN KINASE RECEPTOR ECK PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AI249967
VERSION   AI249967.1 GI:3846496
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 493)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Seq primer: -40UP from Gibco
            unknown library type
            High quality sequence stop: 297.
            Average insert size 1.72 kb. Life Technologies catalog #:
FEATURES   Location/Qualifiers
     source   1..493
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2004483"
                     /clone_lib="NCI_CGAP_Pan1"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
BASE COUNT   107 a   141 c   142 g   101 t   2 others
ORIGIN

```
Query Match      13.8%;   Score 112.2;  DB 44;  Length 493;
Best Local Similarity   57.7%;  Pred. No. 1.6e-23;
Matches 218; Conservative  0; Mismatches 159; Indels  1; Gaps  1;

Qy   71 cgcagcctccaatgaagtcaatctactggattcaaaacaattcaagggagctggggct 130
         || ||  || ||||| |||||| ||||||||||||||||||||||||||| |||||
```

RESULT 10

```
Db   12 CGGGCCAGGGCCAAGGAAGTGGTACTGCTGGACTTTGCTCAGCTGGAGGGGAGCGTCGGCT 71

Qy  131 ggatcttatccatccacatgggtggaagagatcagtggtgtggatgaacattacac 190
        ||||||||||||||||||||| |||||||||||||  | |||||| ||||||||||
Db   72 GGCTCACACACCCGTATGGCAAAGGGTGGGACCTGATGCAGAGAACATCATGAATGACATGC 131

Qy  191 ccatcaaggacttaccgtgtgtcaattgtcatggaccacgtcaaaacaattgggtggtgagaa 250
        |||||||| || | |||| ||||  | ||  | | |||| ||  ||||||||| || | ||||
Db  132 CGATCTACANGTACTCCGTGTCNCAACGTGATGTCTCGGCAGACCAGGACAACATGGCTCCGCA 191

Qy  251 caaactgggtcccaggaactcagctcagaagagattttatgtgggagctcagttcactctac 310
        ||||||||||| | |||| | |||| || ||| |||| |||| ||||| | |||| | |||||
Db  192 CCAACTGGGTGTACCGAGGAGAGAGCCCTANCCTTCATTGAGCCTANGTTCANGGTTACTGTAC 251

Qy  311 gagactgcaatagacattccattggtgttttaggaacttgcaaggaggagacattcaaacctgtact 370
        | |||||||| |||| ||||| ||| || |||| |  |  |||| ||||||||| |||| |||||||
Db  252 GTGACTGCAACAGCCTTCCCTGGGCGGCCCAGCTCCTGCAAGGAGGAGACTNTGCAACCTTCTACT 311

Qy  371 a-catggagctgtgatgaatcatcatgggtgaaatttcgggagcatcagtttacaaagatt 429
        | |||| |||  |||||  |   |  ||| | |||||||| |||||||||||| ||||||||
Db  312 ATGCCCGAGTCGGACCTGGGACTACGGCACCAAACTTTCAGAAGCGCCTGTTCACCCAAGAATT 371

Qy  430 gacacacattgcaagctgat 447
        ||||| |||||| ||| |||
Db  372 GACACCATTGCGCCCGAT 389
```

LOCUS   AA024123   366 bp   mRNA   EST   21-JAN-1997
DEFINITION m101c02.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:459170 5' similar to gb:U07634 Mus musculus
            receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
ACCESSION   AA024123
VERSION   AA024123.1 GI:1489009
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 366)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On May 8, 1995 this sequence version replaced gi:800409.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276058
FEATURES   Location/Qualifiers
     source   1..366
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:459170"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Seq primer: -28M13 rev2 from Amersham.

FEATURES
 source 1..348
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="IMAGE:459170"
  /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
  /sex="unknown"
  /tissue_type="placenta"
  /dev_stage="adult"

High quality sequence stop: 1.
Location/Qualifiers

1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'
TGTTACCAATCTGAAGTGGGAGGGGCGCGCGGAAATTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT   87 a   94 c   93 g   74 t
ORIGIN

Query Match   Best Local Similarity 58.1%;   Score 103.2;   DB 44;   Length 348;
Matches 200; Conservative 0; Mismatches 143; Indels 1; Gaps 1;   Pred. No. 8e-21;

Qy   98 tggattcaaaacaattcaagggagggagctgggctgggatcctcttatccatcacatgggtggg 157
Db   4 TGGACTTCGTAGCAATGAGGGGGAGAGCTTCGGCGTCGGCCACCTTTATGGCAAAGGGT 63

Qy 158 aagagatcaagtggtgtggatgaacattacacacccatccaggacttaccaggtgtgcaatg 217
Db   64 GGGACCTGATCCATAACATTATGGACGACACATGCCTATCTACACATGTACTCGGTGTGCAAACG 123

Qy 218 tcatgtgaccacagtcaaaacaattggctgagacaaacctgggtccccaggaactcag-ct 276
Db 124 TGGTATCCGGGACCAGTACAACTGGCGTCCCGCACCAACTGGTGTGTACCGGCGAGGAGGCCT 183

Qy 277 cagaagaagattatgtggagctcaagttcacctctacagaagactgcaatagcatcacattggtt 336
Db 184 GAGCGGCATCTCTATTGAGGCTCAAGTTCACGAGCAGAGCTGTAAACAGCTTTCTCGAAGGC 243

Qy 337 ttaggaacttgcaaggagagacattcaacctgtactacatgggagtctgatgatcatgg 396
Db 244 GCCAGCGTCATGCATAGAGACCTTCAACCTCTACTATGCAGAGACAGATGTGGACTATGGC 303

Qy 397 gtgaaatttcgagagcatcagttacaaagatggcaccattgc 440
Db 304 ACCAACTTCCAGAAGGCCCAGTTCACCAAGATTGACACCATCGC 347

RESULT 12
AI337308
LOCUS   AI337308   496 bp   mRNA   EST   18-MAR-1999
DEFINITION tb90b01.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:2062345 3',
  similar to SW:EPB2_HUMAN P29323 EPHRIN TYPE-B RECEPTOR 2 PRECURSOR
  ;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION   AI337308
VERSION   AI337308.1   GI:4074235
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 496)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)

BASE COUNT   69 a   103 c   102 g   92 t
ORIGIN

Query Match   Best Local Similarity 63.3%;   Score 110.2;   DB 27;   Length 366;
Matches 169; Conservative 0; Mismatches 98; Indels 0; Gaps 0;   Pred. No. 5.9e-23;

Qy 185 acacaccatccaggacttaccaggtgtgcaatgtcatggacacagtcaaaacaattggc 244
Db 290 ACATGCCTAACTACATGACTACTCGGTCTGTGCAAACGTGGTATCCGGGACCAGGACAACTGCG 231

Qy 245 tgagaacaaactgggtccccaggaactcagctcagaagaagatttatgtgggagctcaagttca 304
Db 230 TCCGCACCAACTGGGTGTGTACCGGCGAGGAGGCCGACCGCAGCTCTTATTGACCTCAAGTTCA 171

Qy 305 ctctacaggactgcaatagcatccaattggtttaggaacttgcaaggagagacattcaacc 364
Db 170 CGGGTGCGAGACTGTAGACAGCCTTCCCGGGGTGGGCGGCTCATGCAATCGAAAAGAGACCTTCAACC 111

Qy 365 tgtactacatggagctgatgatcatgggggtgaaatttcgagagcatcagttacaa 424
Db 110 TCTACTATGCAGAGATGTGGACTATGGGGACTAACTTCCAGAAGGCCCAGTTCACCA 51

Qy 425 agattgaccaccatcgcacgtgatgataaa 451
Db   50 AGATTGACCACCATCGCCCCCTGACGACGAGA 24

RESULT 11
AI323334
LOCUS   AI323334   348 bp   mRNA   EST   23-DEC-1998
DEFINITION mi0lc02.x1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
  clone IMAGE:459170 3', similar to gb:M59371 TYROSINE-PROTEIN KINASE
  RECEPTOR ECK PRECURSOR (HUMAN); gb:U07634 Mus musculus
  receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
ACCESSION   AI323334
VERSION   AI323334.1   GI:4057763
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 348)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Jan 17, 1998 this sequence version replaced gi:2043939.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276058
            This clone was previously sequenced on the 3' end only, this new
            data is from the 3' end
            Trace considered overall poor quality

RESULT 13
LOCUS   AW045268   314 bp   mRNA   EST   18-SEP-1999
DEFINITION UI-M-BH1-akn-b-03-0-UI.s1 NIH_BMAP_M_S2 Mus musculus cDNA clone

Left record

COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 756   Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 479.

FEATURES   Location/Qualifiers
     source   1. .496
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2062345"
                     /clone_lib="NCI_CGAP_Col6"
                     /tissue_type="colon tumor, RER+"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Col0 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneids 1057416-1061255, and 1144584-1145351).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT   104 a   141 c   148 g   103 t
ORIGIN

Query Match
Best Local Similarity   55.9%;   Score 92.6;   DB 45;   Length 496;
Matches 204; Conservative 0; Mismatches 149; Indels 12; Gaps 1;   Pred. No. 1.6e-17;

```
Qy  229 agtcaaaacaaattggctgagaacaaactgggtcccaggaactcagctcagaagatttat 288
         |||| || |||||| |||| |||| ||||  |  |||  || |||  |||  |||||  ||
Db   11 AGCCAGAAACAAACTGGCTACGGACCAAGTTATTCCGGCGCCGTGGCGCCCACCGCATCCAC 70

Qy  289 gtgagctcaagttcactctacagagactgcaatagacattccattggtttaggaactgc 348
         ||| |  ||||||  | |  |  | |||||| || ||  ||||| || ||| || ||||
Db   71 GTGGAGATGAAGTTTTTCGGTTCGGTGCGTGACTGCAGCAATCCCCAGCCGTGCCTGGCTCCTGC 130

Qy  349 aaggagacattcaacctgtactacatggagtctgtgatgaatcatg---------gg 396
         ||||||||| ||||||| ||| || |||| |||| || ||| | ||         ||
Db  131 AAGGAGACCTTCAACCTCTATATACTATGAGGCTGACTTTGACTCGGCCACCAAGACCTTC 190

Qy  397 gttgaaatttcgaagacatcagttacaaaagattgacacccattgcagctgatgaaagtttc 456
         | |    ||| |||| | || ||||  |  || |||| |ccatt||| |||| || |||||
Db  191 CCCAACTGGATGAGGAATCCATGGGGTGAAGGTGGATACCATTGCACCCGACGAGAGCCTTC 250

Qy  457 actcaaaatggatcttgggaccgtattctgaagctcaacactgagattagaagtaggt 516
         | |||  |tggatc||ggg | gt|||| ||||| |||| | ||| || ||| || |
Db  251 TCCCAGGTGGAICCTGGGTGGCGCGTCATGAAAAATCAACACCGAGGTGCGGAGCTTCGGA 310

Qy  517 cctgtcaacaagaagggattttattttggcatttcaagatgttggtgcttgtgttgcctg 576
         ||||| | | | ||| | || ||| |||||| ||| | |  | || || ||||||| |||
Db  311 CCTGTGTCCCGCACGCGCTTCTACCTGGCCTTCTACTGTAACGGGGACGGGCGAGTGGCTG 370

Qy  577 gtgtc 581
         || |
Db  371 GTGCC 375
```

Right record

ACCESSION   AW045268
VERSION   AW045268.1 GI:5905797
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 314)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3187490.
            Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mESt@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NIH_BMAP cDNA clones will be made
            available by the means that is soon to be determined. When NIH
            determines the means for distribution of the BMAP cDNA clones, this
            record will be updated accordingly when that means is determined.
            Seq primer: M13 Forward
            POLYA=NO.

FEATURES   Location/Qualifiers
     source   1. .314
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH1-akn-b-03-0-UI"
                     /clone_lib="NIH_BMAP_M_S2"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S2 library is a subtracted library derived from
                     NIH_BMAP_M_S1, which in turn is a subtracted library
                     derived from a mixture of normalized libraries from ten
                     regions of the mouse brain (cerebellum, brain stems,
                     olfactory bulbs, hypothalamus, cortex, amygdala, basal
                     ganglia, pineal gland, striatum, hippocampus). The driver
                     used for subtraction consisted of a pool of 5,000 clones
                     from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
                     obtained from non-normalized and normalized mouse brain
                     spinal cord libraries.
                     TAG_LIB=NIH_BMAP_M_S2
                     TAG_SEQ=CATGG"
                     TAG_TISSUE=olfactory-bulbs

BASE COUNT   72 a   84 c   96 g   61 t   1 others
ORIGIN

Query Match
Best Local Similarity   56.6%;   Score 89;   DB 64;   Length 314;
Matches 164; Conservative 0; Mismatches 126; Indels 0; Gaps 0;   Pred. No. 1.7e-16;

```
Qy   82 aatgaagtcaattctactggattcaaaaacaattcaaggggagctggggatctcttat 141
          | |||||| |  | |  |||  ||  | |  ||||  ||| |||||| |||| || ||
Db   21 AAGGAAGTTGTTTTGTTGGGACTTCCGAGCAATGAAGGAGAGAACCTCGGCGTCAGGCAC 80

Qy  142 ccatcacatgggtgggaagagaatcagttggttgtggatgaacattaccacaccccatcaggact 201
         ||| | | |ggg|ggga | |||| ||| | | ||| ||||||| || | |||  | ||
Db   81 CCCTATGGCAAAGGGTGGGACCTGATCCAGAACATCATGGACGACACGATCCTATCTACATG 140

Qy  202 taccaggtgtgcaatgtcatggaccacagtcaaaacaaattggctgagaacaaactgggtc 261
         || | ||tgtgcaa  |   || ||  || | |  | |||||| | | |||| ||||||t
Db  141 TACTCGGTGTGCAAACGTGGTATCCGGCGCCGACCAGGACAACTGGCTCCGCCACCAACTGGGTG 200
```

```
Qy    262 cccaggaacctcagcttagagagattatgtgggagctcaagttcaacttacgagggactgcaat 321
          ||| || |||||| ||| |||||| ||||||||||||| ||||| || |||||||| ||||
Db    201 TACCGGGAGGAGGGCCGAGCGCAATCTTTATTGAGCTCAAGTTCACGGTGCGAGACTGTAAC 260

Qy    322 agcattccattggttattaggagaacttgcaaggagagacattcaacctgtacta 371
          ||||||||| |||| ||| |||||| ||||| |||||||||||||| |||||||
Db    261 AGCTTCCCGGGTGGCGGCCAGCTCATCGCANAGAGACCCTTCAACCTCTACTA 310
```

RESULT 14
LOCUS   AI967980/c
DEFINITION wu1lf04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2516671 3'
            similar to SW:EPA8_MOUSE O09127 EPHRIN TYPE-A RECEPTOR 8 PRECURSOR
            ;, mRNA sequence.
ACCESSION   AI967980
VERSION   AI967980.1 GI:5764798
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 418)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3189146.
            Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tel: (301) 496-1550
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
FEATURES   Location/Qualifiers
     source   1..418
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2516671"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469004-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT   73 a   152 c   116 g   77 t
ORIGIN
```

Query Match        10.8%;    Score 88.2;    DB 63;    Length 418;
Best Local Similarity   62.9%;    Pred. No. 3.3e-16;
Matches  154;    Conservative    0;    Mismatches  88;    Indels   3;    Gaps   1;

```
Qy    553 gatgtgtggtgttgcttggttgttctgtgagagtatacttcaaaaagtgccatt 612
          |||| ||| | ||||||||||||| || ||||||||| |||||| |||||||||||
Db    417 GACATAGGTGCCTGCCTCGCCTCTCTCCGCATCTACTATAAGCAAGTGCCCCTGCC 358

Qy    613 acagtgaagaactcggctatgttccagacacggtaccccagtggactcccagtcc---ctg 669
```

```
Db    357 ATCGTGCCGCCAAATCTGGCCTGCCTTCTCCGGAGCAGTGACGGGGGCCCGAGACTCGTCGTCACTG
          ||| |||| | ||| || || | || ||| ||||| ||||| ||| ||| |||||||||| |||
Qy    670 gtggaggttaggggtgtcttgtgtcaacaaattctaaggaggagaagatcctccaaggatgtac 729
          ||||||||||||||||| |||||||||||||| ||||||||| | |||||||||||||||||||||||
Db    297 GTGGAGGTGAGGGGGCCAGTGCCGCACTCAGAGACAGAGGACGCACACCCAAGAGATGTAC 238

Qy    730 tggcagtacagaggcaggtggaatggcgttgtaccacttggcaagtgttcctgaatggtggctat 789
          | |||||||||||||| || ||||| || ||||||||| ||| ||||||||||||| |||||||||
Db    237 TGGCAGCGCGGGGGAGGGGGAGTGTGGGAGCTTCGTGCCCATCGACAAATGCCGTGTGCGGCCGTGGTAC 178

Qy    790 gaaga 794
          |||-|
Db    177 GAGGA 173
```

RESULT 15
LOCUS   AW366937/c
DEFINITION IL0-HT0156-251099-132-a12 HT0156 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW366937
VERSION   AW366937.1 GI:6871587
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 344)
AUTHORS   HCGP http://www.ludwig.org.br/ORESTES/
TITLE   The FAPESP/LICR Human Cancer Genome Project
JOURNAL   Unpublished (1999)
COMMENT   On Mar 16, 1998 this sequence version replaced gi:2961824.
            Contact: Simpson A.J.G.
            Ludwig Institute for Cancer Research
            Laboratory of Cancer Genetics
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HT0156-
            251099-132-a12&t3=1999-10-256t4=1)
            High quality sequence start: 30
            High quality sequence stop: 343.
FEATURES   Location/Qualifiers
     source   1..344
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0156"
                     /clone="IL0-HT0156"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research) into
                     the pUC 18 vector. Reverse transcription of
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
            Seq primer: puc 18 forward
BASE COUNT   73 a   98 c   95 g   78 t
ORIGIN
```

Query Match   10.6%;   Score 86;   DB 80;   Length 344;
Best Local Similarity   60.1%;   Pred. No. 1.4e-15;
Matches 143;   Conservative   0;   Mismatches 95;   Indels   0;   Gaps   0;

```
Qy    182 attacacacccatccaggactaccagtgtgtgcaatgtcatggaccacaagtcaaaacaatt 241
          || || ||||| | | ||| |||||||| ||| || || ||||| || |||| ||||||
Db    268 ATGACATGCCGATCTACARGTACTACCGGTGTCAACGTGACATGGTGTCGGCCACCAGGACAACT 209

Qy    242 ggctgagaacaaactgggtcccaggaggaactcagcagctcagagaagatttatgtgggagctcaagact 301
```

```
Db  208 GGCGCCGCCACCAACTGGGTGTGTACCGAGGAGGAGGCTGAGCGTATCTTCATTGAGCTCAAGT
            |||  ||| |||||||||      |       ||| ||     ||| |  ||    ||||||||||

Qy  302 tcacttctaacgaagactgcaatagcattccattggtttttaggaacttgcaaggagacattca 361
            |  |||| |||| ||| ||  || || |      |  |||||||||| ||| |||||

Db  148 TTACTGTACGTGACTGCAACAGCTTCCCTGGTGGCGGCCAGCTCCTGCAAGGAGACTTTCA 89
            |||||||||||| ||||       ||||| |  |||||||||      |||||||||||||  ||||

Qy  362 acctgtactacatggagtctgatgatgatcatggggtgaaatttcgagagcatcagtt 419
            ||||  |||||     |  |||||       |    || |  |||   | |   |||

Db  88  ACCTATACTATGCCGAGTCAGGACCTGGACTACGGCTCCAATTCCAGAAGCCCTGTT 31
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run On: May 15, 2000, 10:51:21 ; Search time 102.62 Seconds
(without alignments)
1005.931 Million cell updates/sec

Title: US-09-104-340-5
Perfect score: 813
Sequence: 1 atgggattgtcagctctccat.........aaagaggtttatgtgccaa 813

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 226296 seqs, 63486255 residues

Total number of hits satisfying chosen parameters: 452592

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6_COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Accession	Description
1	813	100.0	813	3132	5	US-08-167-919A-9	Sequence 9, Appl1
2	813	100.0	813	3132	5	US-08-715-106-9	Sequence 9, Appl1
3	545	67.0		3254	1	US-08-162-809-15	Sequence 15, Appl
4	351.4	43.2		3906	1	US-08-469-537A-102	Sequence 102, App
5	351.4	43.2		4165	2	US-08-442-248-1	Sequence 1, Appl1
6	351.4	43.2		4165	2	US-08-440-815-1	Sequence 1, Appl1
7	349.8	43.0		4322	2	US-08-673-789-1	Sequence 1, Appl1
8	348.2	42.8		3592	3	US-08-469-537A-100	Sequence 100, App
9	344.4	42.4		3116	4	US-08-469-537A-14	Sequence 14, App
10	344.4	42.4		3116	4	US-08-449-645A-14	Sequence 14, Appl
11	344.4	42.4		3116	6	US-08-702-367A-14	Sequence 14, Appl
12	344.4	42.4		3116	6	PCT-US95-04681-14	Sequence 14, Appl
13	344.4	42.4		3348	1	US-08-222-616-34	Sequence 34, Appl
14	344.4	42.4		4529	4	PCT-US95-04228-34	Sequence 34, Appl
15	344.4	42.4		4529	4	US-08-449-645A-16	Sequence 16, Appl
16	344.4	42.4		4529	4	US-08-702-367A-16	Sequence 16, Appl
17	343.4	42.2		3116	6	PCT-US95-04681-16	Sequence 16, Appl
18	343.4	42.2		3162	4	US-08-449-645A-12	Sequence 12, Appl
19	343.4	42.2		3162	4	US-08-702-367A-12	Sequence 12, Appl
20	334.8	41.2		3162	6	PCT-US95-04681-12	Sequence 12, Appl
21	334.8	41.2		2323	5	PCT-US96-00419-6	Sequence 6, Appl1
22	334.8	41.2		2901	6	PCT-US96-00419-4	Sequence 4, Appl1
23	245.2	30.2		4304	6	PCT-US96-00419-1	Sequence 1, Appl1
24	245.2	30.2		3056	1	US-08-162-809-21	Sequence 21, Appl
25	245.2	30.2		3059	1	US-08-162-809-3	Sequence 3, Appl1
26	219.6	27.0		3125	1	US-08-162-809-19	Sequence 19, Appl
27	213.4	26.2		2820	1	US-08-162-809-5	Sequence 5, Appl1
28	213.4	26.2		4049	1	US-08-162-809-17	Sequence 17, Appl
29	194.4	23.9		4097	1	US-08-162-809-11	Sequence 11, Appl
30	194.4	23.9		2962	4	US-08-449-645A-10	Sequence 10, Appl
31	194.4	23.9		2962	4	US-08-702-367A-10	Sequence 10, Appl
32	150	18.5		2962	6	PCT-US95-04681-10	Sequence 10, Appl
33	150	18.5		3546	1	US-08-162-809-9	Sequence 9, Appl1
34	141.4	17.4		3591	1	US-08-162-809-13	Sequence 13, Appl
35	136	16.7		3776	1	US-08-162-809-7	Sequence 7, Appl1
36	136	16.7		2982	1	US-08-348-143-2	Sequence 2, Appl1
37	136	16.7		2982	2	US-08-571-785-2	Sequence 2, Appl1
38	136	16.7		4027	1	US-08-348-143-3	Sequence 3, Appl1
39	136	16.7		4027	2	US-08-571-785-3	Sequence 3, Appl1
40	136	16.7		4027	2	US-08-162-809-1	Sequence 1, Appl1
41	115.4	14.2		4027	1	US-08-348-143-4	Sequence 4, Appl1
42	80.8	9.9		3133	2	US-08-571-785-4	Sequence 4, Appl1
43	80.8	9.9		3969	1	US-08-436-044-5	Sequence 5, Appl1
44	80.8	9.9		3969	1	US-08-222-616-23	Sequence 23, Appl
45	80.8	9.9		3969	3	US-08-436-054-5	Sequence 5, Appl1
	80.8	9.9		3969	6	PCT-US95-04228-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-167-919A-9
Sequence 9, Application US/0816791919A
Patent No. 5674691
GENERAL INFORMATION:
  APPLICANT: Boyd, Andrew W.
  APPLICANT: Simpson, Richard J.
  APPLICANT: Wicks, Ian
  APPLICANT: Ward, Larry D.
  APPLICANT: Wilkinson, David
  TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
  TITLE OF INVENTION: AND USE THEREOF
  NUMBER OF SEQUENCES: 14
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Scully, Scott, Murphy & Presser
    STREET: 400 Garden City Plaza
    CITY: Garden City
    STATE: New York
    COUNTRY: U.S.A.
    ZIP: 11530
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/167,919A
    FILING DATE: 18-APR-1994
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 435
    FILING DATE: 21-JUN-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PK6841 (AU)
    FILING DATE: 21-JUN-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PK9992 (AU)
    FILING DATE: 12-DEC-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/AU92/00294
    FILING DATE: 19-JUN-1992
  ATTORNEY/AGENT INFORMATION:
    NAME: Digiglio, Frank S.
    REGISTRATION NUMBER: 31,346
    REFERENCE/DOCKET NUMBER: 9159
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (516) 742-4343
    TELEFAX: (516) 742-4366
    TELEX: 230 901 SANS UR
  INFORMATION FOR SEQ ID NO: 9:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 3132 base pairs
      TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 100..3048
US-08-167-919A-9

Query Match                100.0%;  Score 813;  DB 1;  Length 3132;
Best Local Similarity      100.0%;  Pred. No. 2.9e-253;
Matches  813;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 atggattgtcaagctctccatcctcctccttcctcagctgtctgttctctcgacagcttcggg 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 ATGGATTGTCAGCTCTCCATCCTCCTCCTTCCTTCTCAGCTGTCTGTTCTCTCGACAGCTTCGGG 159

Qy    61 gaactgaattccgccagccttccacatgaagtcaatcctactggattcaaaaaacaattcaaggg 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 GAACTGAATTCCGCCAGCCTTCCAACTCCTCCTCCTTCCTTCTCAGCTGTCTGTTCTCTCGACAGCTTCGGG 219

Qy   121 gagctgggctggatctcttatcatcacatgggtggggaagagatcaagtggttgtggatgaa 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 GAGCTGGGCTGGGATCTCTTATCACATCACATGGGTGGGGAAGAGATCAGTGGTGTGGGATGAA 279

Qy   181 cattacacaccctaccaggtgttgcaatgtcatgggaccacagtcaaaacaat 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 CATTACACACCATCAGGACTTACCAGGTGTGCAAATGTCATGGGACCACAGTCAAAACAAT 339

Qy   241 tggctgagaacaaccatgggtcccagaggaactcagctccgaagagagttttatgtggagctcaag 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340 TGGCTGAGAACAAACTGGGTCCCAGGAACTCAGCTCAGAAGATTTATGTGGAGCTCAAG 399

Qy   301 ttcaactctacgagactgcaatagcatccattccattggtttaggaacttgcaaggaggacaattc 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   400 TTCACTCTACGAGACTGCAATAGCATCCATTCCATTGGTTTAGGAACTTGCGAAGGAGGACAATTC 459

Qy   361 aaactgtactacatggagagctgtgatgatggggtgaaatttcagagcatcagttt 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   460 AACCTGTACTACATGGAGTCTGATGATGATCATGGGGTGAAAATTTCGAGAGCATCAGTTT 519

Qy   421 acaaagattgacaccattgcacgtctgatgaaagtttcactcaaatgatcttgggggaccgt 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   520 ACAAAGATTGACACCATTGCAGCGTCTGATGAAAGTTTCACTCAAATGATCTTGGGACCGT 579

Qy   481 atcctgaagctcaacactgagattagagaagtaggtcctgtcaacaagaaggagattttat 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   580 ATCTGAAGCTCAACACTGAGATTAGAGAAGTAGGTCCTGTCAACAAGAAGGAGATTTTAT 639

Qy   541 ttggcattcaagatgttggttgcttgttgtcttgtgagagtactacttcaaa 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   640 TTGGCATTCAAGATGTTGGTTGCTTGTTGTCTTGGGTGTCCTTGAGAGTATACTTCAAA 699

Qy   601 aagtgcccatttacagtgaagaatctgctatgtcttcagacacagggtacccatggaactcc 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   700 AAGTGCCCATTTACAGTGAAGAATCTGGCTATGTTTCCAGACACGGGTACCCATGGGACTCC 759

Qy   661 cagtcccctggttgaagttagagggggtcttgtgtcaacaattctaaggaggaagatcctccca 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   760 CAGTCCCCTGGTGGAGGTTGAGAGGGGTCTTGTGTCAACAATTCTAAGGAGGAAGATCCTCCA 819

Qy   721 aggatgtactgcagtacagaaggcgaataggccttgtaccacattggcaagtgttcctgcaat 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   820 AGGATGTACTGCAGTACAGAAGGCGAATGGCCTTGTACCACATTGGCAAGTGTTCCTACCAAT 879

Qy   781 gctctggctgtgaagaaaagaggtttttatgtgccaa 813
         |||||||||||||||||||||||||||||||||||||
Db   880 GCTGGCTATGAAGAAAAGAGGAGGTTTTTATGTGCCAA 912
```

```
RESULT 2
US-08-715-106-9
; Sequence 9, Application US/08715106
```

Patent No. 6020306

Query Match   100.0%; Score 813; DB 5; Length 3132;
Best Local Similarity   100.0%; Pred. No. 2.9e-253;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 atggattgtcaagctctccatcctcctccttcctcagctgtctgttctctcgacagcttcggg 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 ATGGATTGTCAGCTCTCCATCCTCCTCCTTCCTTCTCAGCTGTCTGTTCTCTCGACAGCTTCGGG 159

Qy    61 gaactgaattccgccagccttccacatgaagtcaatcctactggattcaaaaaacaattcaaggg 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 GAACTGAATTCCGCCAGCCTTCCAACTCCTCCTCCTTCCTTCTCAGCTGTCTGTTCTCTCGACAGCTTCGGG 219

Qy   121 gagctgggctggatctcttatcatcacatgggtggggaagagatcaagtggttgtggatgaa 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 GAGCTGGGCTGGGATCTCTTATCACATCACATGGGTGGGGAAGAGATCAGTGGTGTGGGATGAA 279
```

```
Qy   181  cattacacacaccaaggacttaccaggtgtgcaatgtcatggaccacagtcaaaacaat  240
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   280  CATTACACACCCATCAGGACTTACCAGGTGTGCAATGTCATGGACCACAGTCAAAACAAT  339

Qy   241  tggctgagaacaaactgggtcccaggaactcagaagagatttatgtggagctcaag     300
          ||||||||||||||||||||||||||||||||||        |||||||||||||||
Db   340  TGGCTGAGAACAAACTGGGTCCCAGGAACTCAGCTCAGAAGAGATTTATGTGGAGCTCAAG 399

Qy   301  ttcacctctacagagactgcaatagcattccattggttcttaggaaactgccaaggacattc 360
          ||||||||||||  |||| |||| ||||||||||||||||||||||||||||||||||
Db   400  TTCACTCTACAGAGACTGCAATAGCATTCCATTGGTTTTAGGAACTTGCAAGGACATTC    459

Qy   361  aacctgtactacattggaagtctgatcatgtgggtgaaattcgagagagcatcagtt     420
          ||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||
Db   460  AACCTGTACTACATGGGAAGTCTGATCATGTGGGTGAAATTCGAGAGAGCCATCAGTTT   519

Qy   421  acaaagattgacacaccattgcagctgatgaaagttcactcaaatggatcttgggaccgt  480
          ||||||||| |||||||||||||||||||||||| |||||||||||||||||  ||||||
Db   520  ACAAAGATTGACACCATTGCAGCTGATGAAAGTTTCACTCAAATGGATCTTGGGGACCGT   579

Qy   481  atttctgaagctcaacaactggattagaagaagtcctgtcacaacagaagaggggatttat 540
          |||||| ||||| ||||||||||| ||||||||||||| |||||||||||||||||||||
Db   580  ATTTCTGAAGCTCAACACTGAGATTAGAGAACTAGGTCCTGTCAACAACAGGAGAGGGATTTAT 639

Qy   541  ttggcattcaaagatgttggtgctctgtgtctctgtgagagtatacttcaaa           600
          ||||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db   640  TTGGCATTTCAAGAGATTTGTGGTGCTCTGTGTCTCTGTGAGAGTATACTTCAAA        699

Qy   601  aagtgcccatttacagtgaagaatctggctatgttcccagacacggtgtaccccatggactcc 660
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   700  AAGTGCCCATTTACACGTGAAGAAAATCTGGCTATGTTCCAGACACGGTGTACCCCATGGACTCC 759

Qy   661  cagtccctggtggaggttagagggtcttgtgtcaacaaatctaagaggaagaatcctcca    720
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   760  CAGTCCCTGGTGGAGGTTAGAGGGTCTTGTGTCAACAATTCTAAGGAGGAAGATCCTCCA    819

Qy   721  aggatgtactcgcagtcagaaagcgcaatggcgttgttgtgtcacaagtgttcctgcaat   780
          |||||||||| |||||||||||||||||||||||||||||||||| |||||||||||
Db   820  AGGAGTACTGCAGTCAGAAAGCGCAATGGCTTGTTGTGTCACCATTGGCAAGTGTTCCTGCAAT  879

Qy   781  gctggctatgaagaagaaggaggttttatgtgccaa  813
          |||||||||||||||||||||||||||||||||||
Db   880  GCTGGCTATGAAGAAGAAGGAGGTTTTAAGTGCCAA  912
```

Query Match 67.0%; Score 545; DB 1; Length 3254;
Best Local Similarity 80.0%; Pred. No. 1.4e-166;
Matches 641; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

```
Qy    13  ctctcaccaatcctcctcctctcagctgctctgttctgttctcggggaactgattccg   72
          |   |   |      |  |   | |   |  || |||| ||| | ||| |  |  |
Db    41  CGCCGCGTCGCGCTGCTGCTCTCTCGCGCTCCCCTCGCGCTCCGGCTCCGGCGTCGAGCGCC 100

Qy    73  cagcctccaatgaagtcaatctactggattcaaaaacaattcaaggggaggctgggtgg  132
          | | |  || |||||  || | || ||| |||  | |||||||||  | || |  | ||
Db   101  CGCCCCGGCCAACGAAGTTAAATCTGGATTCAAAAACAATTCAAGGCGCGGGGAGGCTGGG  160

Qy   133  atctcttatccatcacatgggttgtggaagaaatcagtcggttgtggatgaacattacacc  192
          | |||| | | |||||||||| | ||||||||||||||||||| | ||||| ||||| |
Db   161  AtCTCCTACCCATCACATCACATGGGTGGGAAGAAATCAGTCGGTTGTGGATGGACATTACACCA 220

Qy   193  atcaaggacttaccaggtgtgcaatgtcatggaccacaagtcaaaacaattgggctgaaaca 252
          || || |||||| |||||| |||||  |||||||  | |||| ||| |  || || | | |
Db   221  ATCAGAACTTACCAGAGAGCAATGTTATGGACCACAGTGGTGTTGATGGACATTATACTTCCA 280

Qy   253  aactgggtccccaggaactcagctacagaaagatttatttgggacttcaagttctctacga  312
          ||||| |||| ||||||| ||| | || || | |||||| |||| ||| |||| |||| |
Db   281  AACTGGGATTCCACGCACAATTCACCGCAGAAGATATATGTGGACGGACTTCAAAGTTTACCTGAGG 340

Qy   313  gactgcaatagcattccattggttttaggaacttgcaaggagacattcaacctgtactac  372
          | |||||||| ||||| |||||   || ||| | | |  ||| ||| | || | | | |
Db   341  GACTGCAATAGTATCCCTCTACTTCTCTGGGCACTTGCAAAGAGACTTTCAATTGTTGTTAC  400

Qy   373  atggagtctgatgatcatggggttgaaattcgagagagcatcagtttacaaagattgac  432
          |||| |||||| | || |||||| ||| ||||| ||||| ||||||| ||| |||||
Db   401  ATGGGAATCCGATGATGATGACCATTGGCAAAGTTCAGGAGCACCAATTTACGAAGAGTTGAC  460

Qy   433  accattgcagctgataaagttcactcaaatggatcttgggaccgtatctttgaagctc  492
          |||||||| |||| |||||  ||| ||||||||| ||||||||||||||| ||||||
Db   461  ACCAATGGCGGGTCTGATGAGAGCTTCACCCAGATGGAATCTTGGGGACACGGGGATTCTCAAACCTG 520

Qy   493  aacactgagattagagaagtaggtcctgtcaacaacagaagaggatttattggcatttcaa 552
          ||||| |||| ||||| |||| | |||||||||||||| ||||| ||| |||| ||||
Db   521  AATACCGAAAGTCCGCGAGGTGGCACCGTTAGTAAGAAGAGGCGTTTTACTTGGCTTTTCCAA 580

Qy   553  gatgttgggtgcttgttgtttgccattggtctgtgtggagagtatacttcaaagtgccaattt 612
          ||||||||| |||||||||| || |||||| |||||||||| ||| |||| |||| ||||
Db   581  GATGTAGGTGCAATGTTTGCCATTGGGTTGCCTTAGTCTTGCAGGGTGTGACTTCCATGAGAAGTGCCCTTTC 640

Qy   613  acagtgaagaatctggctatgtttccagacacggtaccccagtccctggtg  672
          |||||||||||| || |||| ||| || |||||||||  |||||||||||
Db   641  ACTGTCAAGAACCTCGGCCATGTTTCCAGATACAGTTCCTATGGACTCCCAGTCCCTGGTG 700

Qy   673  gaggttagagggtcttgtgtcaacattctaaggaagaaatcctccaaggagatgatgactgc 732
          |||||| | ||||||||| |||||||  ||| || || ||||| ||||||||| ||||
Db   701  GAGGTGCGCGGGGTTCTTGTGTCAATCAATTCAAAGGAAGAGGACCACCCAAGAGTGACTGC 760

Qy   733  agtacagaaggcgaatgggcttgtacccaatggcaaagtgtctggggctatgaa  792
          |   | |  |  | ||  ||  || |||||| | | | |||||| |||| |||
Db   761  AGCACGGAAAGGAGAATGGCTAGTGCTGCCCATAGGAAGTGCTTGTTGTGTCACCATATGGCTATGAA 820
```

```
Qy    793  gaagagggtttttatgtgccaa  813
           ||||| |||||||  ||||||||
Db    821  GAGAGAGGCTTTGCGTGCCAA  841
```

RESULT   4
US-08-469-537A-102
; Sequence 102, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
;   APPLICANT: Maisonpierre, et al.
;   TITLE OF INVENTION: EHK AND ROR TYROSINE
;   TITLE OF INVENTION: KINASES
;   NUMBER OF SEQUENCES: 107
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Regeneron Pharmaceuticals, Inc.
;   STREET: 777 Old Saw Mill River Road
;   CITY: Tarrytown
;   STATE: NY
;   COUNTRY: U.S.A.
;   ZIP: 10591
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/469,537A
;   FILING DATE: 06-JUN-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 08/406,247
;   FILING DATE: 17-MAR-1995
;   APPLICATION NUMBER: USSN 08/144,992
;   FILING DATE: 28-Oct-1993
;   APPLICATION NUMBER: USSN 07/736,559
;   FILING DATE: 26-JUL-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kepler, Ph.D., Gail M
;   REGISTRATION NUMBER: 32,143
;   REFERENCE/DOCKET NUMBER: REG 070C
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 914-345-7400
;   TELEFAX: 914-345-7721
;   TELEX:
; INFORMATION FOR SEQ ID NO: 102:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3906 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 476..3493
; US-08-469-537A-102

Query Match   43.28; Score 351.4; DB 3; Length 3906;
Best Local Similarity   67.38; Pred. No. 6.7e-104;
Matches 512; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

```
Qy     56  tcggggaactcgagccttccaatgaagtcaatctactggattcaaaaacaattc  115
           ||||| |||| ||| |  ||||||||| |||| || |  ||||| |||||||||||
Db    630  TCCGGACCCTTTGGCCAGCCCCAGTAACGAAGTGAATTATTGGATTCGCGCACTGTCC  689

Qy    116  aaggggagctggcctggatcctcattccatcacaatcagcagtggtggtgtgg  175
           ||  |||||| || ||||||| |  || ||||||||| |||||||||  ||||||
Db    690  TGGGAGAACCTTGGATGGAGTTGCTTTTCCAAAGAAGAATTGGGTTGGAAGTTG  749

Qy    176  atgaacattacacacccatcaggacttaccccggttgtgcaatgtgtcatggggaccacagtcaa  235
           ||||| |||||||| |  ||  ||||||||||||| || ||||| || ||| ||  ||||  |
```

RESULT   5
US-08-442-248-1
; Sequence 1, Application US/08442248
; Patent No. 5759863
; GENERAL INFORMATION:
;   APPLICANT: Caras, Ingrid W.
;   APPLICANT: Winslow, John W.
;   TITLE OF INVENTION: AL-1 Neurotrophic Factor
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/442,248
;   FILING DATE: 15-MAY-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/330128
;   FILING DATE: 27-OCT-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Torchia, Timothy E.

```
Db    750  ATGAAAACTATGCCCCATCCACCTATCACCTATCAAGTGTCCAAAGTTATGGAACAGAATCAGA  809

Qy    236  acaattggctgagagaacaaactgggtccccagggaacctcagctcagaagaagagattatgtggagc  295
           ||||| ||  ||  ||||| |||  |||| ||| || || ||| || ||  | |||  |  |||  |||
Db    810  ATAATTGGCTGTTGACCAGTTGATCTCTTAACGAAGGTGCTTCCAGAATTTTTATTGAAC  869

Qy    296  tcaagttcacttcacgagactgcaatagcatccaatggtcttttaggaacttgcaaggaga  355
           |||||||| || |||||  ||||| ||||| |||| ||| |||||  || ||||  |||||
Db    870  TCAAGTTCACTCTGAGGGGATTGCAACAGCCTTCCTCTGGAGGACGTGGGGGACTTGCAAGGAGA  929

Qy    356  cattcaacctgtactacatggagtctgatgatcatgggtgtgaaatttcgagagcatc  415
           || |||||||||||| |||| ||||||||| || |||||||| ||  ||  |||||||
Db    930  CCTTTAACATGTATTATTATTTTGACTCGGATGATGAGGAATATCAAAGAGAACC  989

Qy    416  agtttacaaagttgacaccattgcagctgatcgaaagtttcactcaaatgatcttgggg  475
           || ||||||||||  ||| |||| |||| ||  || |||||||| ||||| ||  |||
Db    990  AGTACATCAAGAGCCGATACCATTGCTCTGATCGAAGCTTCACGACAGACTTGACCTTGGAG  1049

Qy    476  accggtattctgaagctcaaacactgagattagagaagtaggttcctgtcaacagaaggat  535
           || ||  || |||||||||| |||| || |||||||||| || ||||||| ||| ||||||
Db   1050  ACCGGGTCATGAAGCTGAAATACGGAGGTCAGAGATGTAGGAGGTCCTGAGCAAAAGGGAT  1109

Qy    536  tttatttggccattcaagatgttcggttcgttgttgccttggttgttctgtgagagtatact  595
           || |||  || |||||||||  |||| ||  ||||| |||  ||||||||  || ||| |||
Db   1110  TTTATCTTGCTTTCCAACGAATGCGGTGCTTTGTCGTCCGTGTCCTACT  1169

Qy    596  tcaaaaagtgcccattacagtgaagaaatctggctatgtttcccagctacaccatgg  655
           || ||||||||| || ||||  ||||| ||||| |||||| || ||  |||| |||
Db   1170  ATAAAAAGTGCTTCTTTCTGTAGTTAGACAHTTGGCTGTTTTCCCTGACACGATCACTCGGAG  1229

Qy    656  actcccagtccc---ttggtgaggttgagggtctgttgtcaacaattctaaggaggaag  712
           |||                  || ||||||  |||  || ||  || |||  ||||||
Db   1230  CAGATTCTTCCCAGTTCCTAGGAGTGTCCAGGGTCTCCGTCAACCAHTCTGTGACGACG  1289

Qy    713  atcctcccaaggatgtactgcagtacagaaggcggaatgggtggcttcggtttcccatgttcaagtgtt  772
           ||| || ||||  ||| ||  ||  ||||||||| ||| ||  ||| ||| | ||||||| || |||||
Db   1290  ATCCTCCCAAAATGCAATTGCACATTGCAGTGCTCAAAGGGGAGTGGCTGGTTCCCATCGGGGAAATGCA  1349

Qy    773  cctgcaatgctggctatgaagagaaagggtttttatgtgccaa  813
           |||                   ||  |||||||||||||||||
Db   1350  TGTCCAAGGCCGGATATGAAGAGACAAAAATGGGTACCTGTCAA  1390
```

RESULT   6

; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 920C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4165 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-442-248-1

Query Match   43.2%; Score 351.4; DB 2; Length 4165;
Best Local Similarity   67.3%; Pred. No. 6.9e-104;
Matches 512; Conservative   0; Mismatches 246; Indels   3; Gaps   1;

Qy   56 tcggggaactgattccgcagccttccaatgaagtcaatctactggattcaaaaacaattc   115
           ||||| | | ||||| |||| ||| | ||| | || ||||||| | |||| |||
Db   695 TCCGGACCCTTTTGGCCAGCCCAGTAACGAAGTGAATTTATTGGATTCCGCGCACTGTCC   754

Qy   116 aaggggagctgggctggaatctcttatccatcacatgggtggaagagagatcagtggtgtgg   175
           | ||| || |||| ||||| || | ||| |||||| | ||| || || ||| |||| |||
Db   755 TGGGAGACCTTGGAtGGAATTGCTTTTCCAAAGAAATGGGTGGGAGAGAGAATTGGTGAAGTTG   814

Qy   176 atgaacattacacaccaccaggacttaccaggtgtgcaatgtcatgtaccacagtcaaa   235
           || | | | ||| ||||| ||| | || || | | ||| || | || ||||
Db   815 ATGAAAACTATCCCCCATCCACACTATCAAGTGTGCAAAGTTATGGAACAGAAATCAGA   874

Qy   236 acaattggctgagaacaaactgggtccccaggaactcagctcagaagagatttatgtggagc   295
           ||| ||| || | | | |||| ||||||| ||| ||| |||||| || || | |||||
Db   875 ATAATTGGCTGTTGAACGCCTTCCCTGGACACCTTACAGAAGCCTTGCAACGCCTT   934

Qy   296 tcaagttcactctacgagactgcaatagacattccattggttttaggaaacttgcaaggaga   355
           || ||||| ||| | || ||| || ||| ||| || | ||| |||||| ||| ||| |||
Db   935 TCAAGTTTACTCTGAGGGATTGCAACAGCCTTCCTGGAGAACTGGGACTTGCAAGGAGA   994

Qy   356 cattcaacctgtactacatggagtctgatgatgatcatggggtgaaattcgagagcatc   415
           ||||| | || | || |||| | ||| ||| |||||||| |||| | ||| | |||||
Db   995 CCTTTAACATGTATTATTTTGAGTCGGATGATGAGGAGAAATATCAAAGAGAACC   1054

Qy   416 agtttacaaagattgacaccaattgcagctgatgaaagtttcactcaaatggatcttgggg   475
           |||| | ||| || || |||||| | ||| |||| | || ||| ||| || ||||
Db   1055 AGTACATCAAGATCAAGATCGATACCATTCGTCGCTGATGAGAGCTTCACCGAACTTGGACCTTGGAG   1114

Qy   476 accgtattctgaagctcaacactgagattagagaagtaggctccttgtcaacaagagaaggat   535
           || | |||| |||| ||| |||||| ||||||| ||| || |||| | |||| ||||
Db   1115 ACCGGGTCAtGAAGCTGAATACGGAGAGGTCAGAGAGGTAGGACCTCTGAGCAAAAAGGGAt   1174

Qy   536 tttatttgcattcaaagtgttggttgcttgccttgtttgtgagagatatact   595
           ||| |||||| | | |||| | || | || ||| |||||| |||| |||
Db   1175 TTTATCTTGCTTCCCAAGATGTCGGTtGCTTGCATCGGCTTGTTTCGTGTCCGTGTCTACT   1234

Qy   596 tcaaaaagtcgcccattacagtgaagaatctggctatgttcccagacacggtacccatgg   655
           || ||| | || ||| || |||| ||| || | |||| ||||| | ||| | |||
Db   1235 ATAAAAAAGTCCTTCTGTTAGTGTTTAGACATTGGCTGTGTTTCCCTGACACGATCACTGGAG   1294

Qy   656 acttccaagtccc---tggtggaggttagagggtcttgtgtcaacaattctaaggaggaag   712
           || | | | ||| |||| |||| |||| | | |||| | || ||| |||||
Db   1295 CAGAtTCTTCCCAGtTGCTCTAGAGGtGTCAGGCTCTGCGGTCAACCATTCGTGTGACAGACG   1354

Qy   713 atccctccaaggatgtactccagtacagaggacggaatgggctcttgtgtaccccaattggctcaagtgtt   772
           | || ||||| | |||| ||||||||||| | ||| || ||| | | || |||| ||||| |
Db   1355 ATCCTCCAAAATGCAATTGCAATTGCTGCTGAAGGGAGtGGGCTGGCTGGTGCTTCCCATCGGGAAAATGCA   1414

Qy   773 cctgcaatgctggctcatgaagaagaagaaggtttatgtgccaa   813
           ||| || ||| |||||| ||||| || || || | ||| ||
Db   1415 TGTGCAAGGCCGGATATGAAGAGAGAAAAAATGGTACCTGTCAA   1455

; US-08-440-815-1
; Sequence 1, Application US/08440815
; Patent No. 5798448
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; APPLICANT: Winslow, John W.
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,815
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330128
; FILING DATE: 27-Oct-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 920C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4165 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-440-815-1

Query Match   43.2%; Score 351.4; DB 2; Length 4165;
Best Local Similarity   67.3%; Pred. No. 6.9e-104;
Matches 512; Conservative   0; Mismatches 246; Indels   3; Gaps   1;

Qy   56 tcggggaactgattccgcagccttccaatgaagtcaatctactggattcaaaaacaattc   115
           ||||| | | ||||| |||| ||| | ||| | || ||||||| | |||| |||
Db   695 TCCGGACCCTTTTGGCCAGCCCAGTAACGAAGTGAATTTATTGGATTCCGCGCACTGTCC   754

Qy   116 aaggggagctgggctggaatctcttatccatcacatgggtggaagagagatcagtggtgtgg   175
           | ||| || |||| ||||| || | ||| |||||| | ||| || || ||| |||| |||
Db   755 TGGGAGACCTTGGAtGGAATTGCTTTTCCAAAGAAATGGGTGGGAGAGAGAATTGGTGAAGTTG   814

Qy   176 atgaacattacacaccaccaggacttaccaggtgtgcaatgtcatgtaccacagtcaaa   235
           || | | | ||| ||||| ||| | || || | | ||| || | || ||||
Db   815 ATGAAAACTATCCCCCATCCACACTATCAAGTGTGCAAAGTTATGGAACAGAAATCAGA   874

Qy   236 acaattggctgagaacaaactgggtccccaggaactcagctcagaagagatttatgtggagc   295
           ||| ||| || | | | |||| ||||||| ||| ||| |||||| || || | |||||
Db   875 ATAATTGGCTGTTGAACGCCTTCCCTGGACACCTTACAGAAGCCTTGCAACGCCTT   934

Qy   296 tcaagttcactctacgagactgcaatagacattccattggttttaggaaacttgcaaggaga   355
           || ||||| ||| | || ||| || ||| ||| || | ||| |||||| ||| ||| |||
Db   935 TCAAGTTTACTCTGAGGGATTGCAACAGCCTTCCTGGAGAACTGGGACTTGCAAGGAGA   994

Qy   356 cattcaacctgtactacatggagtctgatgatgatcatggggtgaaattcgagagcatc   415
           ||||| | || | || |||| | ||| ||| |||||||| |||| | ||| | |||||
Db   995 CCTTTAACATGTATTATTTTGAGTCGGATGATGAGGAGAAATATCAAAGAGAACC   1054

Qy   416 agtttacaaagattgacaccaattgcagctgatgaaagtttcactcaaatggatcttgggg   475

RESULT 7

US-08-673-789-1
Sequence 1, Application US/08673789
Patent No. 5814479

```
Db 1055 AGTACATCAAGATACGATACCAATTGCTCCTGATGAGAGACCTTCACCGAACTTGACCTTGGAG
            |||  ||||||    |||   |||||   |||||| ||||  ||||||| |||||||||||||||
Qy  476 accgtatctctgaagctccaacctgagcttagcagaggtaggtcctgtcaacaagaggggat 535
Db 1115 ACCGGGTCATGAAGCTGAATACGGAGGTCAGAGATGTAGGACCTCTGAGCAAAAAGGGAT 1174
            |||||||  ||||   |||||||||||||||||  |||||||||| ||||||||||||
Qy  536 tttatttggcattcaagaatgttggtcgcttgtgtctgtctgtgagagtatact 595
Db 1175 TTTACTGCTTCAAGAATGTCGGGTGCTTGCCATCGCTCTGGTTTTCCTGGTGTCCGTGTCTACT 595
            ||||  |  ||||||||||| |||| |||| |   |||||||||  ||||| ||||| ||||
Qy  596 tcaaaaagtgcccattacagtgaagaaatctggctatgttcccagacacgtaccatgg 655
Db 1235 ATAAAAAATGTCCTTCTGTAGTTAGACATTTGGCTGTTTTCCCTGACACGATCACTGGAG 1234
            ||| ||| ||| ||  || |  |||   |||| ||| |||||| |||||| ||| ||||
Qy  656 actcccagtccc---tggtggaggttagagggtctttgtcaacaattctaaggaggaag 712
Db 1295 CAGATTCTTCCCAGTTGTCTAGAGGTGTCAGCCTCCTCGCGTCAAACCAATTCTGTGACAGACG 1294
            |   ||| ||||   |||| ||||| ||  ||| ||    |||||| |||||| |||||| ||
Qy  713 atcctccaaggatgtactgcagtacagaaggcgaatggcgtatgttgtaccccattggcaagtgtt 772
Db 1355 ATCCTCCAAAATGCATTGCAGTCGTCGAAAGGGAGTGGCTGGTTCCCATCGGGAAATGCA 1354
            |||||||||| || ||||||||  | |||||| |||||| |||| ||||| |||||||||
Qy  773 cctgcaatgtggctatgaagaagaaggtttatgtgccaa 813
Db 1415 TGTGCAAGGCCCGGATATGAAGAAGAGAGAAAAATGGGTACCTGTCAA 1455
```

GENERAL INFORMATION:
  APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
  APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
  APPLICANT: GEORGE, F.
  TITLE OF INVENTION: BSK RECEPTOR LIKE
  TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
  TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
  TITLE OF INVENTION: METHODS
  NUMBER OF SEQUENCES: 14
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: MORGAN & FINNEGAN
    STREET: 345 PARK AVENUE
    CITY: NEW YORK
    STATE: NEW YORK
    COUNTRY: USA
    ZIP: 10154
  COMPUTER READABLE FORM:
    MEDIUM TYPE: FLOPPY DISK
    COMPUTER: IBM PC COMPATIBLE
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WORDPERFECT 5.1
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/673,789
    FILING DATE:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 435
    CLASSIFICATION: 435
    FILING DATE:
    APPLICATION NUMBER: 08/177,812
    FILING DATE: 04-JAN-1994
  ATTORNEY/AGENT INFORMATION:
    NAME: CAROL M. GRUPPI
    REGISTRATION NUMBER: 37,341
    REFERENCE/DOCKET NUMBER: 2026-4105
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (212) 758-4800
    TELEFAX: (212) 751-6849
    TELEX: 421792
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 4322
    TYPE: NUCLEOTIDE
    STRANDEDNESS: DOUBLE

TOPOLOGY: UNKNOWN
US-08-673-789-1

Query Match 43.0%; Score 349.8; DB 2; Length 4322;
Best Local Similarity 67.1%; Pred. No. 2.3e-103;
Matches 511; Conservative 0; Mismatches 247; Indels 3; Gaps 1;

```
Qy   56 tcggggaactgatccgcagcctccaatgaagtcaatctactggattcaaaacaattc 115
            ||||   ||  || ||||||  ||||  ||||| ||| |||| |||||   || |
Db  572 TCCGGACCCTTTGCCCAGCCCACGAACCAACGAAGTGAATTTGCCGCCACTGTCA 631
Qy  116 aaggggagctgggcctggatctcttttatccatcacatgggtgggaagagatcagtgggtgg 175
            |||||||||||||  ||||||||  ||||| ||||||||||||||||||||| |||| |||
Db  632 TGGGGGACCTTGGATGGAATTGCTTTTTCCAAAGAACACGGGTGGGAAGAGGATTGGGAAGTTG 691
Qy  176 atgaacattacacacccatcaggaacttaccaaggtgttcaagtcatggacacagtcaaa 235
            |||| | ||||| ||| ||| || |||| |||||| || ||||| ||||| ||||| |||
Db  692 ATGAGAACTATGCCCCATCCACACAATACCAAGTGTGCAAAGTTATGGAACAGAAATCAGA 751
Qy  236 acaaattggctgagaacaacaactgggtcccaggagaacctcagcagaagagcttatgttggagc 295
            || |||||||| || ||||||  ||| ||||| || ||||||| |||||||  |||| ||| ||
Db  752 ATAATTGGCTGGTTGACCAGTTGGATCTCTAACGAACGGTGCTTCCAGAATTCTTTATTGAAC 811
Qy  296 tcaagttcaactctacgagagactccaaatagcaattccaatggtttttaggaacttgcaaggaga 355
            ||||||| | || ||| || |||||||  || ||||||||||| |||||||||||| |||||||||
Db  812 TCAAGTTTACTTTAAGGACTGCCAACAGCCCTTCCTGGAGGACTGGGGGACTTGTTAAGGAGA 871
Qy  356 cattcaaacctgtactacatgggagtctgatgatgtatcatggggtgaatttcgagagcatc 415
            |||| ||| |||||||||||||||| |||||||||||||||||| |||||||| ||||| |||
Db  872 CATTTAACATGATTATTTGAATCAGATGAGATGAGAAGTATCAAAGAGAACC 931
Qy  416 agtttacaaagatttgacaccattgcgctgatgaaagtttcactcaaatggatcttgggg 475
            |||||||||||||||||||| ||| |||||||| |||||||||| ||||||||  ||| ||
Db  932 AATACATCAAGATTGATGATACCAATCGGTCTCGCAGATGAGAGGCTTCAGAGAACTTGANCTTGGTG 991
Qy  476 accgtatctgaagctcaacactgagattagagaagtaggtcctgtcaacaagaggggat 535
            ||||| |||||||||||  ||||| ||||| ||||||||||||| |||||||||||||||
Db  992 ACCGTGTCATGAAGAACTGAAATACAGAGGTCAGAGATGTAGAGAGATTGGGTAGAAGGGAT 1051
Qy  536 tttatttggcattcaagaatgttggtcgcttgtgtctgtctgtgagagtatact 595
            |||||||| |||||||||||||||||| |||||| |||| ||||||||||| |||
Db 1052 TTTATCTTGCTTCCAAGAATGTCGGGTGCTTGCGATTGCTCTGGTTTTCGTGGTGTCCTACT 1111
Qy  596 tcaaaaagtgcccattacagtgaagaatctggctatgtctataccaggacacggtacc---ca 652
            ||||||||||| |||||||||| ||||| |||||||||  |||| ||||||| |||||   ||
Db 1112 ATAAAAAAGTGTCCCTCTGTAGTAGAGACAACTTGGCTATTTCCCTGACACTATCACTCGGAG 1171
Qy  653 tggactccccagtccctggtgggaggttgagggggtcttgtgtcaacaattctaaggaggag 712
            |  ||| ||||||||||||||||||||| || ||| || |||||||||||||||||||||||
Db 1172 CAGATTCATCACAGTTGTTAGAGGTGTCAGCCTCCTGCGTCAACCATTCTGTGACAGATG 1231
Qy  713 atcctccaaggatgtactgcagtacagaaggcgaatggcgtatgttgtacccattggcaagtgtt 772
            |||||||| |||||||||||||| | |||||| ||||||||||||| |||||| |||||||||||
Db 1232 ATCCTCCCAAGATGCATTGCAGTGCTGAAAGGGAGTGGCTGGTTCCCATTGGGAAATGCA 1291
Qy  773 cctgcaatgtggctatgaagaagaaggtttatgtgccaa 813
            |||||||||||||||||||||||||| || ||||||||||
Db 1292 TGTGCAAGGCTGGATATGAAGAAGACAAAAATGGGTACCTGCCAA 1332
```

RESULT 8

US-08-469-537A-100
Sequence 100, Application US/08469537A
Patent No. 5843749

GENERAL INFORMATION:
  APPLICANT: Maisonpierre, et al.
  TITLE OF INVENTION: EHK AND ROR TYROSINE
  TITLE OF INVENTION: KINASES
  NUMBER OF SEQUENCES: 107
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Regeneron Pharmaceuticals, Inc.
    STREET: 777 Old Saw Mill River Road

```
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette
    COMPUTER: IBM Compatible
    OPERATING SYSTEM: DOS
    SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/469,537A
    FILING DATE: 06-JUN-1995
    CLASSIFICATION: 435
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: USSN 08/406,247
    FILING DATE: 17-MAR-1995
    APPLICATION NUMBER: USSN 08/144,992
    FILING DATE: 28-OCT-1993
    APPLICATION NUMBER: USSN 07/736,559
    FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
    NAME: Kempler, Ph.D., Gail M
    REGISTRATION NUMBER: 32,143
    REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 914-345-7400
    TELEFAX: 914-345-7721
    TELEX:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
    LENGTH: 3592 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: unknown
    TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
    NAME/KEY: CDS
    LOCATION: 598..3444
FEATURE:
    NAME/KEY: modified_base
    LOCATION: 56
    OTHER INFORMATION: /mod_base= OTHER
    OTHER INFORMATION: /label= N
    OTHER INFORMATION: /note= "Where N = G, A, C or T"
FEATURE:
    NAME/KEY: modified_base
    LOCATION: 3538
    OTHER INFORMATION: /mod_base= OTHER
    OTHER INFORMATION: /label= N
    OTHER INFORMATION: /note= "Where N = G, A, C or T"
US-08-469-537A-100
```

```
Query Match     42.8%;  Score 348.2;  DB 3;  Length 3592;
Best Local Similarity   68.0%;  Pred. No. 7e-103;
Matches  501;  Conservative    0;  Mismatches  233;  Indels    3;  Gaps    1;

Qy   79 tccaatgaagtcaatcactggatcaaaaacaattcaaaggggacgctgggcctggatctct 138
Db  688 TCCAACAAGTGTGTTGCTTGATACATCTACAGTGATGGGAGAACTAGGAGAAAACA 747

Qy  139 tatcatcacatgggtgggaaagatcaagtcggtggatgaacattaccacccatcagg 198
Db  748 TATCCACTGAATGGGTGGGGGATGCCATTACTGAAATGGATGAACACAGGCCCATACAT 807

Qy  199 acttacaggtgtgcaatgtcaatggaccacagtcaaaacaattggctgagaacaactgg 258
Db  808 ACATACAGGTATGCAATGTCATTGGAACCAAAACCAGAACAACTGGCTTCGTTGCTTCGTTG 867

Qy  259 gtcccaggaactcagcgctcaggaagtttatgtggagctcaactcactacgaacctgc 318
Db  868 ATCTCTCGTGATGCTGCTCAGAAAATCTATGTGGAAATGAAGTTCACATTGAGAGATTGT 927
```

```
Qy  319 aatagacattccattgtttttaggaaacttgcaaggaggaaggagacaattcaaacctgtactacatggag 378
Db  928 ACAGCAATCCCARGGGTCTTGGGAACGTGCTAAAGAAACATTTACTCGTATTATATTGAA 987

Qy  379 tctgatgatgatcatgggtgaaatttcgagagcgcatcagttttacaaagattgaccaccatt 438
Db  988 TCTGACCAAATCCCACGGAACTAAAATTCAAGCCAAGCCAATATATATAAAGATIGACACAATT 1047

Qy  439 gcagctgatgaaagtttcaactcaaaatggatctctggggaccgtattctgaagctcaacact 498
Db 1048 GCTGCGGATGAGAGTTTACTCAGATGGATTTGGGGTCAATACCTTAAACTCAACACT 1107

Qy  499 gagattaggaagaggtaggtccctgtcaacagaaagggattttattttgtcattcaaagtgtt 558
Db 1108 GAAGTTCGTGAGGTGGGCGCCAATAGAAAAGGAAAGGATTCTATTGGCTTTTCAAGATATT 1167

Qy  559 ggtgtcttgttgttgccttggtgtctgtgagagtatacttcaaaagtgcccatttacagtg 618
Db 1168 GGAGCAGCAGCATTGCTCTGGGTCTCAGTCCGTGTGTTCTACAAAAAAGCCCCTTCACTGTG 1227

Qy  619 aagaatctggctatgtttccagacacagtgtaccca---tggactcccagtccctgtggag 675
Db 1228 CGGAACTTGGCTATGTTTCCGGATACCATCCCAAGGGTTGACTCTTCCTTGCTTGAA 1287

Qy  676 gttagagggtgtcttgtcaacaaattctaaggaggaggaagatcctccaagggatgtactgcagt 735
Db 1288 GTCCGGGGGCTCATGCGTGAAGAGTTCTGAGGAGCGAGATACTCCTAAACTCTACTGTGGA 1347

Qy  736 acagaagagcgaatggcttgtacccattggcaagtgttcctgcaatgctggtgctatgatgaagaa 795
Db 1348 GCTGATGGAGACTGGCTCGTCGCTCCCTCTTGGAGGTGTATCTCCACTACAGGGTATGAAGAGA 1407

Qy  796 agagggttttatgtgcca 812
Db 1408 ATCGAGGGTTCTTGCCA 1424
```

```
RESULT 9
US-08-449-645A-14
Sequence 14, Application US/0849645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3116 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
```

```
;
;                   NAME/KEY: CDS
US-08-449-645A-14   LOCATION: 34..2994

Query Match
Best Local Similarity   42.4%;    Score 344.4;   DB 4;   Length 3116;
Matches 509;  Conservative  0;  Pred. No. 1.1e-101;  Mismatches 231;  Indels  6;  Gaps  2;

Qy   74 agccttccaatgaagtcaatctactggattcaaaacacaattcaaaggggggagctgggctgga 133
        ||||   | ||||| | ||||| || | |  |||| ||| || |||||||| |||  || ||
Db  110 ACCCCGGAATGAAGTTACCTTATTGGAATTCCAGATCTGTTCAGGGAGGAGAACTTGGGTGGA 169

Qy  134 tctc---ttatccatcacatgggtgggaagagatcagtggtgtggatgaacattacacac 190
        |  |    | ||| || |||  |||| |||  || || || |||| | ||| |||| ||||
Db  170 TAGCAAGCCCTCTGGAAGGAGGAGGTGAGTATCATGGATGAACCAGCTACATGGTACAC 229

Qy  191 ccatcaaggacttacacagttgtgcaatgtcatgtggaccacagtcaaaacaattgggctgagaa 250
        || |||| | || |  | | || | | | | |||| | ||| || |||| || ||  | ||
Db  230 CAATCCGAACCTACCAAGTGTCCAAATGTGATGGAACCCAGCCAGAATAACTGGCTACGAA 289

Qy  251 ccaaactggggtcccaggaactcagctcagagaagattatgtggagctcaagttcactctac 310
        | || ||| |  | || | | ||  | ||| || || || |||| |||| ||  || ||
Db  290 CTGATTGGATCACCCGAGAAGGGCGTCANGGGACTCGAAGGACGAGTTTAACCTGTACT 349

Qy  311 gagactgcaatagcattccattggttttaggaacttgcaaaggacgcattcaacctgtact 370
        | |||| || | ||  | ||| || || || | ||| || ||||| |||| || ||
Db  350 GGGACTGCAATAGTCTTCCGGGGCGTCATGGGGGACTTGCAAGGAGACGTTTAACCTGTACT 409

Qy  371 acatggagtctgatgatcatgggtgggtgaaatttcgaaggcatcaaacctggact 430
        | |||| ||| |  |||| || ||| | ||||  | |   ||||| |||| || |||
Db  410 ACTATGAATCAGACAACGACAAAGAGCGTTTCATCAGAGAGAACCAGTTTGTCAAAATTG 469

Qy  431 acaccattgcagctgatgaaagttcactcaaatggatctctgggggaccgtattcttgaagc 490
        |||| |||||| ||| | || |||  | | ||||| |||| | || ||| ||| || ||||
Db  470 ACACCANTTGCTGCTGANGANGAGCTTCACCCAGTGGACNNCGTGTTCTAAAAGGTGAGAGC 529

Qy  491 tcaacctgagattagaagaagtagctcctgtcaacaagaagaggggattttattggcatttc 550
        |  | || ||| || |  |||  | | | || ||||| || |||| || ||  | ||||||
Db  530 TGAACACCGAGAATCCGGGATGTAGGCCCATTAAGCAAGAAAAAGGGGTTTTTACCNNGCCAT 589

Qy  551 aagatgtgttgtctgtgtcctgtgagagtatactccaaaagtgcccat 610
        || | |||| || | | || | | || || | |  ||  |||| ||||
Db  590 AGGATGTGGGGGGCCTGCAANCCAGTTCCCGTGTCATCCCGTGTCATCCCAC 649

Qy  611 ttacagtgaagaatctggctatgtttccagacacggtaccccatg---gactcccagtccc 667
        | |||||| || ||  | | ||| | || | | ||| |  |   |    || || || |||
Db  650 TCACAGTCCGCAATCTGGCCCAATCTGCTGTATAAAAGTGTCGAC 709

Qy  668 tggtggaggttagaggggtcttgttgtcaacaattctaaggaggagagtcctcaaggatgt 727
        ||| || ||| || | |  | ||| | || ||||| ||  | || |||||| || ||
Db  710 TGGTGGAAGTTCGGAGGCCTCCGTGTCAACAACTCAGAGAGAAAAGGATGTGGCAAAAAGT 769

Qy  728 actgcaagtacgaagagcgaatgggtcttgttgtaccccattggcaagtgttcctgcaatgctggct 787
        || | || |  |  |||  ||| | ||| | |  | || ||| ||| ||  |  | |||| |||
Db  770 ACTGTGGGCAGGCAGATGGTGAATGGCCTGGTACCCATTGCAACGCCTNNGCAACGCTGGGGC 829

Qy  788 atgaagaaagagggtttatgtgccaa 813
        ||||| || |||| | |  || ||||
Db  830 ATGAGGAGCGCGAGCGGGAAATGCCAA 855
```

RESULT 11
PCT-US95-04681-14
; Sequence 14, Application PC/TUS9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..2994

Query Match   42.4%; Score 344.4; DB 6; Length 3116;
Best Local Similarity   68.2%; Pred. No. 1.1e-101;
Matches 509; Conservative   0; Mismatches 231; Indels   6; Gaps   2;

```
Qy     74 agccttccaatgaagtcaatctactggattcaaaaacacattcaagggagctgggctgga 133
          || ||      |||||||||  |||||  ||       |||| |||||||| || |||||
Db    110 ACCCCGGAATGAAGAGTTACCTTATTGGAATTCCAGATCTGTTCGTGTGGATGACATTGGGTGA 169

Qy    134 tctc---ttatccatcacatgggtggggaagagatacagtgtggtgatgaacattacacac 190
          ||||     | ||  |||| |  ||  |||| ||| || ||||||| ||| ||  | || ||
Db    170 TAGCAAGACCCTCTGGAAGGAGGTGGGGAGAAGTGAGTATCATGGAATGAAAAATACAC 229

Qy    191 ccatcaaggacttaccagtgtgtcatgtggaccacagtcaaaacattggctgagaa 250
          | ||||||| | |||  ||||  |||| |||||| ||||| ||||||||  ||| ||
Db    230 CAATCCGAACCTACCAAGTGTGCAATGTGATGGAACCAGCCAGATAACTGGCTACGAA 289

Qy    251 caaactgggtccccaggaactcagtcagaagatttatgtggagctcaagttcactctac 310
          | ||||||| ||||| ||||||| || |  ||| |||||| |||| | | ||||| |||
Db    290 CTGATTGGATCACCCGAGAAGGGGCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGA 349
```

RESULT 12
US-08-222-616-34
; Sequence 34, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goedel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935

FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEX: 910/371-7168
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-616-34

Query Match   42.4%;   Score 344.4; DB 1;   Length 3348;
Best Local Similarity   68.2%;   Pred. No. 1.1e-101;
Matches 509;   Conservative   0; Mismatches 231;   Indels   6;   Gaps   2;

```
Qy   74 agccttccaatgaagtcaatctactggatcaaaaacaattcaaggggagctgggctgga 133
        | ||| ||||||||  |||||| ||| | ||| | |||| ||||||||   |  |||||
Db   77 ACCCCGGAATGAAGTTACCTTATTGGATTCAGATCTGTTCAGGGAGCAACTTGGGTGGA 136

Qy  134 tctc--ttatccatcacatgggtgggaagagatcagtggtgtggatgaacattacacac 190
        ||||    ||| ||   || |||||||| | |||  |||| ||   |||| ||||| |||
Db  137 TAGCAAGCCCTCTGGAGGAGGGCTGGGAGGAAGTGAGTATCATGGATCAAAAAATACAC 196

Qy  191 ccatcaggacttaccaggtgtgcaatgtcatggaccacagtcaaaaacaatggctgagaa 250
        ||  | |  || ||||  | |||||  |||| || ||||| || |||||| |||||
Db  197 CAATCCGAACCTACCAAGTGTGCAAATGTCATGGATCAAAATAACTGGCTACGAA 256

Qy  251 caaactgggtccccaggaactcagctcagaagatttatgtggagctcaagttcacctac 310
        ||||||||||||| ||  |||||| ||| ||||| ||||||||| |||| | ||||||
Db  257 CTGATTGGATCACCCGAGAAGGGGCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGA 316

Qy  311 gagactgcaatagcaattccaattggttttaggaaacttgcaaggagacattcaacctgtact 370
        ||||| |  ||| || ||||  |||| || ||| | || |||| ||||| |||  || ||||
Db  317 GGGACTGCAATAGTCTTCCGGCGGTCATGGGCGACGTTTAACCTGTACT 376

Qy  371 acatggagtctgatgatcatgggtgaaattcgagaagcatcagttacaaagattg 430
        ||| ||| |||| ||   |  || |  |||| |||| |  | |||| ||||| ||||
Db  377 ACTATGAAATCAGACAACAGAACGGTTCATCAGAGAGAACCAGTTTGTCAAAATTG 436

Qy  431 acaccattgcagctgatgaaagtaggtccctgcaaatggatcttgggagaccgtattctgaagc 490
        | ||||||||  ||| |||||| ||||||| || |||||  ||| |||| || | |||||| |||
Db  437 ACACCATTGCTCCTCGATGAGAGCTTCACCCAAGTGGACAATTGGTGACAGAATCATGAAGC 496

Qy  491 tcaaacactgaattagaagaataggtcctgtcaacagaaggggatttttattgctcattc 550
        |||||| ||||  || ||| ||||| |||| || ||| ||| |||||| ||| ||| |||
Db  497 TGAACACCGAGATCCGGGATGGTAGGCCCATTAAGCAAAAAGGGGTTTTTACCTGGCTTTTC 556

Qy  551 aagatgttgtcttggttggcttggtgtctgtgagagtatactttcaaaaagtgcccat 610
        || | ||| |  ||| ||||| |||| || |  | | || ||||| ||| ||| ||||
Db  557 AGGATGTTGGGGGCCTTGCCATCGCCCTGGTATCAGTCCGTGTTCTATAAAAAGTGTCCAC 616

Qy  611 ttacagtgaagaatctggctatgttcccagacacgtacccatg--gactcccagtccc 667
        || ||| ||||| |||| ||| ||||| ||||| |||  || |  ||  || ||||||
Db  617 TCACAGTCCCGCAATCTCGGCCCAGTTTCCTGACACCATCACAGGGCTGTCGAATACGTCTCC 676

Qy  668 tggtgaaggttagagggtctgttgttgtcaacagttctccacaaggatgt 727
        ||| |||||  || |||| || || || ||  || |||| || ||| |||
Db  677 TGGTGGAAGTTCGAGGGCTTCCTGTGTCAACAACTTCAGAGAAGAGTGCCAAAAATGT 736

Qy  728 actgcagtacagaaggcgaatggcttgtacccgaatctcctgcaatgctggct 787
        ||||| |  || || || |||||| |||  | || ||||| ||||| |||||
Db  737 ACTGTGGGGCCAGATGTTGAACTGGGCTTCCTGTCAACAACTGCTCGAAGGACGCTTGGGC 796

Qy  788 atgaagaagaggtttatgtgccaa 813
        || |||||| ||| |||| ||||
Db  797 ATGAGGAGGCGAGCGGAGGCGGAGAATCCAA 822
```

RESULT 13
PCT-US95-04228-34
Sequence 34, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEX: 910/371-7168
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-04228-34

Query Match   42.4%;   Score 344.4; DB 6;   Length 3348;
Best Local Similarity   68.2%;   Pred. No. 1.1e-101;
Matches 509;   Conservative   0; Mismatches 231;   Indels   6;   Gaps   2;

```
Qy   74 agccttccaatgaagtcaatctactggatcaaaaacaattcaaggggagctgggctgga 133
        | ||| ||||||||  |||||| ||| | ||| | |||| ||||||||   |  |||||
Db   77 ACCCCGGAATGAAGTTACCTTATTGGATTCAGATCTGTTCAGGGAGCAACTTGGGTGGA 136

Qy  134 tctc--ttatccatcacatgggtgggaagagatcagtggtgtggatgaacattacacac 190
        ||||    ||| ||   || |||||||| | |||  |||| ||   |||| ||||| |||
Db  137 TAGCAAGCCCTCTGGAGGAGGGCTGGGAGGAAGTGAGTATCATGGATCAAAAAATACAC 196

Qy  191 ccatcaggacttaccaggtgtgcaatgtcatggaccacagtcaaaaacaatggctgagaa 250
        ||  | |  || ||||  | |||||  |||| || ||||| || |||||| |||||
Db  197 CAATCCGAACCTACCAAGTGTGCAAATGTCATGGATCAAAATAACTGGCTACGAA 256

Qy  251 caaactgggtccccaggaactcagctcagaagatttatgtggagctcaagttcacctac 310
        ||||||||||||| ||  |||||| ||| ||||| ||||||||| |||| | ||||||
Db  257 CTGATTGGATCACCCGAGAAGGGGCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGA 316

Qy  311 gagactgcaatagcaattccaattggttttaggaaacttgcaaggagacattcaacctgtact 370
        ||||| |  ||| || ||||  |||| || ||| | || |||| ||||| |||  || ||||
Db  317 GGGACTGCAATAGTCTTCCGGCGGTCATGGGCGACGTTTAACCTGTACT 376
```

RESULT 14
US-08-449-645A-16
Sequence 16, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehaviland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: 16: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182

```
Db   317 GGGACTGCAATAGTAGTCTTCCGGGCGGTCATGGGGACTTGCAAGGAGGACGTTTAACCTGTACT 376
Qy   371 acatggagtctgatcgatcatggggtgaaattcgagagcatcagttacaaagattg        430
Db   377 ACTATGAATCAGACAAACGACAAAGAGCGTTTCATCGAGAGAGACAGAAAATTG          436
Qy   431 acaccaattgcagctgatgaaagtttcactcaaatggatcttgggaccgtattctgaagc     490
Db   437 ACACCATTGCTGCTGATGAGAGAGACTTCACCCAGAGGACATTGGTGACCGTATTCTGAAGC    496
Qy   491 tcaacacttagagattagaagtaggtcctgtcaacaagaagtggatttattttggcatttc    550
Db   497 TGAACACCGAGATCCGGGATGTAGGGCCATTAAGCACAAGGAGGGTTTTACCTGGCTTTTTC    556
Qy   551 aagatgttggttgttgtttgccttggtctgtgagtatacttcaaaaagtgcccat        610
Db   557 AGGANGTGGGGGCCTGCATGGCCCTGGTATCAGTCGCTGTGTTCTATAAAAGTGCCAC       616
Qy   611 ttacagtgaagaatctgtatgttcccagacacggtaccccatg--gactcccagtcc       667
Db   617 TCACAGTCCGCAANTCTGGCCCCAGTTTCCTGCANCCATCACAGGGTCCTGATACGTCTTCCC   676
Qy   668 tggtggaggttagagggtcttgtcgtcaacaattctaaggagggaagatctccaaggatgt    727
Db   677 TGGTGGGAAGTTCGAGGGCTCCTGTGTGTCAACAACTCAGAAGAGAAAAGATGTGCCAAAAATGT 736
Qy   728 actgcagtacagaaggcgaatggctcttgtaccatgggcaagtgttcctgcaatgctggct    787
Db   737 ACTGTGGGGCCAGATGGTGAANGGCTGGTACCATTGGCAACTGCCTATGCAACGCGTGGGC    796
Qy   788 atgaagaaaagaggttttatgtgcaa        813
Db   797 ATGAGGAGCGGGAGCGGGAGAATGCCAA        822
```

US-08-449-645A-16

Query Match   42.4%; Score 344.4; DB 4; Length 4529;
Best Local Similarity   68.1%; Pred. No. 1.3e-101;
Matches 512; Conservative   0; Mismatches 231; Indels   9; Gaps   2;

```
Qy    71 cgcagccttccaatgaagtcaatctactggttgatcaaaacaattcaagggagctgggct   130
Db   265 CGCCAGGCTGCGAAGAAAGTACTTACTGCCTGGAATCTCRAAAGCNCAACANAAACAGAGTGGGAGT   324
Qy   131 ggatctcttattccatcacatggggtggggaaggaatcaagtgtgtggatgaacattacac   190
Db   325 GGATTTCCTCTCCACCCANATGGGGGTGGGGAAGAAATTAGCTGGGTTTGGATGAGGAACTATACCC 384
Qy   191 ccatcaaggacttaccaaggttgtgcaatgtgtcatggaccacagtcaaaacaatggctgagaa   250
Db   385 CGATACGAANCANTACCAACAGGTGTCCCAAGTCANTGGAGCCCAACCAACAACCTGGCTGCGGA   444
Qy   251 caaactgggtccccaggggaactcagagagacctcagaagatttatgtggagctcactctac   310
Db   445 CTAACTGGANTTCCAAAAGGCAANTGCACAAAAGGGATTTTGTRAGAAAGTTCACCCTGA       504
Qy   311 gagactgcaatagcattccattggttttaggagaactgcaaggagacaatcaacctgtact   370
Db   505 GGGATTGTAACAGTCTCTTCCGGGGACTACGTGGGAAACTTGCAAGGAGAACATTTAATTTGTACT   564
Qy   371 acatggagtctgatgatcatgggtgaaatttcgagagcatcaagttacaaagattg       430
Db   565 ATTATGAAACAGACAACTATGACACTGGCAGGAANTATAAGAGAAANACTTCTATGTAAAAATAG   624
Qy   431 acaccaattgcagctgatgaaagtttcactcaaatggatcttgggaccgtattctgaagc     490
Db   625 ACACCATTGCTGCGANTGAAANGTTTAACCCAAGGTGACCTTGGTGAAAGAGAGAGAAGC     684
Qy   491 tcaacactgagattagaagtaggtcctgtcaacaagaaggggattttgattcattc        550
Db   685 TTAACACTGAGGGTGAGAGAGAGATTGGAACCTTTGTGTCCAAAAAGGGAATTCTAATCTTGCCATTTC 744
Qy   551 aagatgtttgctgttgtgccttggtctgtgtctgtgagtatactccaaaaagtgccat    610
Db   745 AGGATGTAGGGCGCCTTGCATACCTTTGGTTTGTCTGTCAAAAGTGTACTACAAGGAAGTGCTGGT 804
Qy   611 ttacagtgaagaatctggcctatgttcccagacacggtacc---catggactcccagtcc    667
Db   805 CCATTATNTGAGAACTTAGCTANTCTTTCCAGATACNACAGTGACTGGCTTCAGAATTTTCCTCTT   864
Qy   668 tggtggaggttagagggtctgtgtcaacaattctaaggga-------ggaagatctccaa   721
Db   865 TAGTCGAGGTTCGAGGGACANCTGTGTCAGCAGTGCAGGAGGAAGAAACGCCCCA       924
Qy   722 ggatgtactgcagtacagaaggcgaatggctcttgtaccatgtgtgcaagtgttcctgcaatg   781
Db   925 GGATGCACTGCAGTTGCAGACAGGAGGAATGGGTTAGTGCCCATTGGGAAAATGTATCTGCAAAG    984
Qy   782 ctggtctatgaagaaaagaggttttatgtgcaa        813
Db   985 CAGGCTACCAGCAAAAAGGAGGAGACACTTGTGAA    1016
```

RESULT 15
US-08-702-367A-16
Sequence 16, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehaviland Drive
CITY: Thousand Oaks

```
;
; STATE:  California
; COUNTRY:  USA
; ZIP:  9320
; COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/702,367A
;   FILING DATE:
; CLASSIFICATION:  435
; ATTORNEY/AGENT INFORMATION:
;   NAME:  Winter, Robert B.
;   REFERENCE/DOCKET NUMBER:  A-287
; INFORMATION FOR SEQ ID NO:  16:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  4529 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  186..3182
US-08-702-367A-16
```

```
Query Match          42.4%;   Score 344.4;  DB 4;   Length 4529;
Best Local Similarity  68.1%;   Pred. No. 1.3e-101;
Matches  512;  Conservative   0;  Mismatches  231;  Indels   9;  Gaps   2;
```

```
Qy     71 cgcagccttccaatgaagtcaatctactggattcaaaaacaattcaagggagctgggct 130
          ||||||| ||  ||||  |||||   ||| ||||| ||| |||  ||   ||| || |
Db    265 CGCAGGCTGCGAAGGAGTACTACTGCTGGATTCTAAAGCACAACAAACAGAGTTGGAGT 324

Qy    131 ggatctcttatccatccacatgggtggaagagatcagtgttggatgaacattaccac 190
          |||||  || ||  ||||||||| ||||||||| ||||||||| |||| |||||||||
Db    325 GGATTCCTCTCCACCCAAATGGGTGGGGAAGAAAATTAGTGGGTTTGGATGAACTATACC 384

Qy    191 ccatcaggacttaccaggtgtgcaaatgtcatggaccacagtcaaaacaattggctgagaa 250
          | ||  || ||||||  |||| ||||| ||||| ||||  |||||||| |||| |||||||
Db    385 CGATACGAACATACCAGGTGTCCCAGTCATCGAGCCCACCAAAACAACTGGCTGCGGGA 444

Qy    251 caaactgggtccccaggaactcagctcagagaagattatgtggagtcaagttcactctac 310
          | ||| || |||||||||||||||| ||||||| ||| |||||| ||||||| ||||||||
Db    445 CTAACTGGATTTCCAAAGGCAATGCACAAAGGGATTTTTTGTAGAATTGAAATTCACCCTGA 504

Qy    311 gagactgcaatagcattccattggttttaggaacttgcaaggagacattcaacctgtact 370
          |  |||||||| ||||| |||||||| ||||  |||||||||| |||||||| ||| |||
Db    505 GGGATTGTAACAGTCTTCCTGAGTACTGGGAACTTGCCAAGGAAACATTTAATTGTGACT 564

Qy    371 acatggagtctgatgatgatcatgggtgaaattcgagagacctcagtttacaaagattg 430
          |||||||| | || ||| ||||||||| ||||||||||||| ||||||||||||||||
Db    565 ATTATGAAACACAGACTATGACACTGGCAGGGAATATAAGAGAAACCTCTATGTAAAAATAG 624

Qy    431 acaccattgcagctgtgtgaaagtttcactcaaatgtctggggaccgtattctgaaagc 490
          ||| |||||||||||| |||| ||| || ||| ||| |||| |||| ||||  |||||||
Db    625 ACACCATTGCTCGAGATGAAAGTTTTACCCAAGGTGACCTTCGTGAAAGAAGATGAAGC 684

Qy    491 tcaacactgagattagagaagtaggtcctgtcaacaagagagagattttattggcatttc 550
          ||||||||| || ||| |||| |||||  ||||  |||  ||||| ||||||||||||||
Db    685 TTAACACTGAGTGAGAGATTGGACCTTGTCCAAAAAGGGGATTCATCTACTCTGCCTTTC 744

Qy    551 aagatgttggtcgttcttggtcgttctgtgagagtatactccaaaaagtgccat 610
          |||||| | ||| ||||||  ||||||| |||||||||| || ||||||||||||
Db    745 AGGATGTAGGGGCCTTGCATAGCTTGGGTTCGTCAAGGTGTACTACAAGAGTGCTGGT 804

Qy    611 ttacagtgaagaatctgggctatgttccagacacgtacc---catggactcccagtcc 667
          |||| ||| |||| |||| || ||||||||| |  |||||   ||||  |||||||||||
Db    805 CCATTATTGAGAACTTAGCTATCTTTCCAGATACAGAGTGACTGGTTCAGAATTTTCCTCTT 864

Qy    668 tggtggaggttagagggtctgtgtcaacaattctaagga------ggaagatcctcaa 721
          |||||||||| ||||||| |||| |||||||| ||||||      |||| ||||||||
Db    865 TAGTCGAGGTTCGAGGGGACATGTCACCAGTGCAGTGGAAGAAAGCCGGAAAAACGACCCCA 924
```

```
Qy    722 ggatgtactgcagtaacaggaagccgaatgggcttgtacccattggcaagtgttcctgcaatg 781
          ||||||  | |||   |||  || |||  ||||| ||| || ||| ||  || | || | |
Db    925 GGATGCACTGCAGTGCAGCAAGGTGGTGCCAGAAGGAGAAGGCGGAAGAAATGTATCTGCAAAG 984

Qy    782 ctggctatgaagaagaggtttttatgtgccaa 813
          ||||||| |||  |||  ||| | ||| ||||
Db    985 CAGGCTACCAGCAAAAGGAGAGACACTGTGAA 1016
```

```
Search completed: May 15, 2000, 12:05:01
Job time: 4420 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 10:52:01 ; Search time 111.14 Seconds
                                          (without alignments)
                                          1830.180 Million cell updates/sec

Title: US-09-104-340-5
Perfect score: 813
Sequence: 813
                 1 atggattgtcaagctctccat.........aaagaggtttatgtgccaa 813

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   N_Geneseq_36:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	813	100.0	3132	1	Q34513	HEK coding sequenc
2	545	67.0	3254	1	Q90659	Eph-related PTK Ce
3	351.4	43.2	3906	1	V70207	Rat receptor tyros
4	351.4	43.2	4165	1	T18893	Rat REK7 cDNA. AL-
5	349.8	43.0	4322	1	V58192	Mouse Bsk receptor
6	348.2	42.8	3592	1	V70208	Rat receptor tyros
7	344.4	42.4	3116	1	T02948	EPH-like receptor
8	344.4	42.4	3348	1	T03100	Protein tyrosine-k
9	344.4	42.4	4529	1	T02949	EPH-like receptor
10	343.4	42.2	3162	1	T02947	EPH-like receptor
11	334.8	41.2	2323	1	T32962	EPH-like receptor
12	334.8	41.2	2901	1	T32961	Mouse developmenta
13	334.8	41.2	4304	1	T32960	Mouse developmenta
14	245.2	30.2	3056	1	Q90662	Eph-related PTK Ce
15	245.2	30.2	3059	1	Q90653	Eph-related PTK Ce
16	245.2	30.2	3125	1	Q90661	Eph-related PTK Ce
17	219.6	27.0	2820	1	Q90654	Eph-related PTK Ce
18	218.8	26.9	4281	1	C53471	elk cDNA. Expressi
19	218	26.8	3105	1	T07308	Receptor tyrosine
20	218	26.8	3105	1	T84528	Mouse Nuk tyrosine
21	213.4	26.2	4049	1	Q90660	Eph-related PTK Ce
22	213.4	26.2	4097	1	Q90657	Eph-related PTK Ce
23	194.4	23.9	2962	1	T02946	EPH-like receptor
24	150	18.5	3546	1	Q90656	Eph-related PTK Ce
25	150	18.5	3591	1	Q90658	Eph-related PTK Ce
26	143.4	17.6	3751	1	062461	Human embryonal ki
27	141.4	17.4	3776	1	Q90655	Eph-related PTK Ce
28	136	16.7	2982	1	Q90971	Protein p140 cDNA
29	136	16.7	4027	1	Q90972	Eph-related PTK Ce
30	136	16.7	4027	1	Q90982	Protein p140 cDNA
31	129.6	15.9	3663	1	T72320	Embryonic stem cel
32	136	16.7	4027	1	Q90972	Protein p140 cDNA
33	81.6	10.0	3969	1	Q49757	pTK gene HpTK5. Ne
34	80.8	9.9	3969	1	T03099	Protein tyrosine-k

ALIGNMENTS

35	80.8	9.9	4290	1	Q92641	Human non-differen
36	80.8	9.9	4290	1	T18394	Receptor type tyro
37	80.8	9.9	4290	1	T42593	Coding sequence fo
38	80.8	9.9	4490	1	T51235	Receptor-type tyro
39	64	7.9	450	1	V90061	EST clone CW960. N
40	64	7.9	3673	1	V33697	Human thymus recep
41	64	7.9	4022	1	V62177	Human receptor typ
42	37.4	4.6	324	1	061342	Human brain Expres
43	32.8	4.0	9927	1	X12997	Enterococcus faeca
44	32	3.9	1893	1	Q95110	Human thyroid horm
45	31.4	3.9	4822	1	V04445	Sequence used in d

RESULT 1
Q34513

ID   Q34513   standard; DNA; 3132 BP.
AC   Q34513;
DT   24-MAY-1993 (first entry)
DE   HEK coding sequence.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-i JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   5'utr   1..99
FT   /*tag= a
FT   cds   100..3051
FT   /*tag= b
FT   3'utr   3052..3132
FT   /*tag= c
FT   signal_peptide 100..159
FT   /*tag= d
FT   5'utr   1723..1795
FT   /*tag= e
FT   misc_feature   1..99
FT   /*tag= e
FT   mat_peptide   160..216
FT   /note= "Potential transmembrane region"
FT   /*tag= f
FT   mat_peptide   2617..2679
FT   /note= "Purified HEK protein #1"
FT   /*tag= g
FT   /note= "Purified HEK protein #2"
PN   WO9300425-A.
PN   07-JAN-1993.
PD   19-JUN-1992; AU0294.
PF   19-JUN-1992; AU0294.
PR   21-JUN-1991; AU-0006841.
PR   12-DEC-1991; AU-0009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   WPI; 93-036373/04.
DR   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence encodes human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-l and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the
CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence   3132 BP; 888 A;   711 C;   759 G;   774 T;

Query Match   100.0%; Score 813; DB 1; Length 3132;
Best Local Similarity   100.0%;
Matches   813; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 atggattgtcagctctccaatcctcctccctccttctcagctgtctctgttcctcgggg   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 ATGGATTGTCAGCTCTCCAATCCTCCTCCCTCCTTCTCAGCTGTCTCTGTTCCTCGGG   159

Qy     61 gaactgattccgcagccgccttccaatcctcctccctccttctcagctgttccgacgcttcggg  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 GAACTGATTCCGCAGCCGCCTTCCAATGAAGTCAATCTACTGGATTCAANAACAATTCAAGGG  219

Qy    121 gagctgggctggatcctcttatccatccacatggggtggaagagatcagtggttgtgtggatgaa  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 GAGCTGGGCTGGATCCTCTTATCCATCACACATGGTGGGAAGAGATCAGTGGGTGTGGATGAA   279

Qy    181 cattacacacaccatcaggaccttaccagagttgtgcaatgtcatggaccacagtcaaaacaat   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 CATTACACACCATCAGGACTTACCAGCGTCCAAATCTCATGGACCACAGTCAAAACAAT       339

Qy    241 tggctgagaacaaacctggggtccccaggaaactcagctccagaagacagattttatgtggagctcaag  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    340 TGGCTGAGAACAAACCTGGGTCCCCAGGAAACTCAGCTCCGGAAGACTCACTTGAAG     399

Qy    301 ttcaactctacgagactgcaatagcattccattggttgttttaggaacttgcaaggagacattc   360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    400 TTCACTCTACCAGACTGCAATAGCATTCCATTGGTTGTTTTAGGAACTTGCAAGGAGACATTC  459

Qy    361 aacctgtactactacggttcgatgatcatggggtgaaatttcgagagcatcagttt        420
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    460 AACCTGTACTACTACGGTTCGATGATCATGGGGTGAAATTTCGAGAGCATCAGTTT    519

Qy    421 acaaagagattgacaccattgcacgctgatgatgaagttcactcaagtttcactcaaatgtgaaaa  480
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    520 ACAAAGAGATTGACACCATTGCACGCTGATGATGAAGTTCACTCAAGTTTCAATTGGATCTTCCA  579

Qy    481 atctctaagctcaaacactggagtattagaagaagtagtcctgtcaacaacaaacattggctgagaaca  540
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    580 ATTCTGAAGCTCAACACTGAGATTACGAAGAATTGGCTGCGAAACA                639

Qy    541 ttggcatttcaagatgtgtggttgtctgttgcctgtgagtatacttcaaa            600
          |||||||||||||||||||||||||||||||||||
Db    640 TTGGCATTTCAAGATGTGTGGTTGTCTGTTGCCTGTGAGTATACTTCAAA            699

Qy    601 aagtgcccaattacacgtgagaaatctgggctatgtgttcccaggacacgtgtacccatggctggactcc   660
          |||||||||||||||||||||||||||||||||||||||||||
Db    700 AAGTGCCCAATTACACGTGAGAAATCTGGGCTATGTTCCCAGGACACGTGTACCCATGGCTGGACTCC  759

Qy    661 cagtccctggttggaggttagagggttacttgtgtctgtctcaacaattctcaaggagggaagaatcctcca  720
          |||||||||||||||||||||||||||||||||||||||||
Db    760 CAGTCCCTGGTTGGAGGTTAGAGGGTTACTTGTGTGTCTCAACAATTCTCAAGGAGGTTATTCTGAAGTC  819

Qy    721 aggaatgtactgcagtacagaaggacgaatggcgcttgtgtcaccaatggcgttccatgatgcctgcaat  780
          |||||||||||||||||||||||||||||||||||||
Db    820 AGGATGTACTGCAGTACAGAAGGACGAATGGCGCTTGTGTCACCAATGGCGTTCCTGCAAT       879

Qy    781 gctggctatgaagaagaagggtttatgtgtgccaa                               813
          ||||||||||||||||||||||||||||||||||
Db    880 GCTGGCTATGAAGAAGAAAGAGAGGTTTTATGTGCCAA                            912
```

```
Query Match     67.0%;  Score 545;  DB 1;  Length 3254;
Best Local Similarity   80.0%;   Pred. No. 3.5e-170;
Matches  641;  Conservative  0;  Mismatches 160;  Indels  0;  Gaps  0;

Qy     13 ctctccaatcctcctccctccttctcagctgtctctgttcctcgggggaactgattccg   72
          |||||||||||||||||||||||||||| ||||||| | | |||| | ||||
Db     41 CGCCGCCGCTGCCGCTGCTTCGCTGCTCTGCCGCTCCGCCTCGGGCGTCTGAGCGCC    100

Qy     73 cagccttccaatgaagtcaatctactggattcaaaaacaattcaaggggagctgggctgg  132
          ||  | ||| |  |  ||    |   |||| ||||| ||| ||||| |||||||
Db    101 CGCCCCCCGCCAACGAAGTTAATCTGCGGATTCAAAAACAATTCAAGGGGAGCTGGGCTGG  160

Qy    133 atctcttatccatccacatggggtgaaatttcgagagcatcagtt                192
          ||||||| ||| |||| ||| ||| | |||||| |||||||||||
Db    161 ATCTCCTACCCATCACACATGGGGTGGGAAGAGATGTTAGTGGTGTGGTATACTCCA   220

Qy    193 atcaaggacttaccagtgtgcaatgtcatggaccacagtcaaaacaataacacacc      252
          ||| |||| || || ||| |||| ||||| | ||||| ||||||||| ||||||
Db    221 ATCAGAACTTACCAAGAGCAATGTTATGGATCACACTCAAAACAATTGGCTGCGAAACA   280

Qy    253 aactgggtccccaggaactcagctcagaagaagatttatgtggagctcacttacga     312
          |||| ||||| ||||| ||||| || || |||||| | |  ||||||| |||||
Db    281 AACTGGGATTCCACGCAATTCAGCGCAGAAGATATATGGAGCTTCAAGTTTACCTTGAGG  340

Qy    313 gactgcaatagcattccattggttgttttaggaacttgcaaggagacattccctggtg   372
          ||||| ||||| || ||||| |||| |||  ||  || | |||| |||||||
Db    341 GACTGCAATAGTATCCCTCTAGTTCTGGGCACTTGCAAAGAGACTTTCAAATCTGTATTAC  400

Qy    373 atggagtctgatgatgatcatggggtgaaatttcgagagcatcaaagattgac        432
          |||||| ||| || ||| ||||||  ||||| | ||||||| |||||||||
Db    401 ATGGAAATCCGATGATGACCATTTGGCAAGTTCAGAGCACCAATTTACGAAGAATTGAC   460

Qy    433 accattgcagctgatgaagttcaccaacaaatggaagtcctgtcaacaacaaacattggcc  492
          |||| ||||| ||||| || |||| || |||||  ||||||| ||||||||||||||
Db    461 ACCATGGCGGCTGATGAGAGCTTCACCAATTTGCCAAGTGGGGACCGAGTCTTCAAGCTG   520

Qy    493 aaactgagagtttagaagtaggtcctgtcaacaacaagggagatttattggcatttcaa    552
          |||| ||||| | |||||| | |||||| ||| ||| ||||||| | |||
Db    521 AATACCGAAGTCCGCGAGGTGGGGACCGTTAGTAACAAGGAGATTCTCAAGCTG        580

Qy    553 gatgttggttgtctgttgcctgtgtctgtgagtatacttcaaaaagtgcccattc      612
          |||| || ||||| || || || ||  || || || || |||||||| ||||
Db    581 GATCTAGGTGCATGTGGTTGCCTTAGCAGTAGTAGTCTGTTTACTTGGCTTTCCA   640

Qy    613 acagtgaagaatctggctatgtttcccaggacacgtgtaccatggactcccagtcctgg    672
          ||||| || |||||  |||| |||| ||| ||| ||  |||||  || ||||
Db    641 ACTGTCAAGAACCTCGCCATGTTTCCAGATATCCTAGGACATCAGTTCCTAGGAAGCTG  700

Qy    673 gaggttagagggttcttgtgtgtcttcaacaatctcaaggaggaagatccctccaaggatgtactg  732
          ||||| ||| || || ||| |||||| || |||| ||||| ||||| || || ||
Db    701 GAGGTGGGGGGGTTCTTGTGTCAATCATTCAAGGAGGAGAGCACCACCCAAGATGTACTG  760

Qy    733 agtacagaaggacgcaatggcttcttcctgcaagtgctggctatgaa               792
          |||| |||| ||| ||||| ||  |||| || ||  | |||||||||
Db    761 AGCACGGAAGGAGACAATGCTAGTGCCCATAGGAAAGTCGTTGGTAATGCTGGCTATGAA  820

Qy    793 gaaagaggtttatgtgtgccaa                                        813
          ||||||||||||||||||||||
Db    821
```

RESULT 3

| | | | | | | | | | | | | | | | | | | | | | | |
Db   821 GAGAGAGGCTTTGCGTGCCAA 841

ID   V70207   standard; DNA; 3906 BP.
AC   V70207;
DT   11-FEB-1999 (first entry)
DE   Receptor tyrosine kinase Ehk-1 encoding DNA.
KW   Rat receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis; ss.
OS   Rattus sp.
FH   Key   Location/Qualifiers
FT   CDS   476..3493
FT   /*tag= a
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995; 469537.
PR   06-JUN-1995; 469537.
PR   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Maslakowski P, Yancopoulos GD;
DR   WPI; 99-044584/04.
DR   P-PSDB; W83147.
PT   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Claim 5; Fig 22; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence encodes rat Ehk-1.
SQ   Sequence 3906 BP; 1097 A; 876 C; 990 G; 943 T;

Query Match   43.28%; Score 351.4; DB 1; Length 3906;
Best Local Similarity   67.3%;
Matches 512; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy   56 tcggggaactggattccgcagccttccaatgaagtcaatctactggattcaaaaacaattc 115
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   630 TCCGGACCCTTTGCCAGCCCCAGTAACGAAGTGAATTTATTGGATTCGCCCACTGTCC 689

Qy   116 aaggggagctgggctggatctcttatccatcacatgggttgggaagaagatcaagtggtgtgg 175
          ||| |||||||||||||||||||| ||||||||| |||||||||| ||| || |||||
Db   690 TGGGACGCCTTGGATGGATGGTTGCTTTTCCAAAGAAATGGGTGGGAAGAGAATTGGTGAAGTTG 749

Qy   176 atgaacattacacaccacctcaggacttacaggtgtgcaatgtcatggaccacagtcaaa 235
          ||||||||||||||||||||||||||||||||||||||| |||||| |||||||||||||
Db   750 ATGAAAACTATGCCCCACCTATCCAAGTGTTGCAAAGTTATGGAACAGAGAATCAGA 809

Qy   236 acaattggctgagaacaaactgggtccccccaggagacctaagagagattattatggagagc 295
          ||||||||||||||||||| |||| ||| |||||||||||||||||||||||||||||||
Db   810 ATAATTGGCTGTGTTGACCAGTTGGATCTCTAACGAAGGTGCTTCCAGAATTTTTATTGAAC 869

Qy   296 tcaagttcactctacgagactgcaatagacattccattggttttaggaaacttgcaaggaga 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 TCAAGTTTACTCTGACGGGATTGCAACAGCCTTCCTGGAGGACTGGGACTTGCAAGGAGA 929

Qy   356 cattcaacctgtactacagtggagtctgatgatggggtgaaattcgagagcatc 415
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 CCTTTAACACTGTATTATTTTGAGTCGGATGAATGGGAGAAAATATCAAAGAGAACC 989

Qy   416 agtttacaaagagattgacacacattgagacgctggtctgatgaaagttcactcaaatgatctgggg 475
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 AGTACATCAAGATCGAATATACCAATTGCTCGTCGATGAGACTTCACCGAGTCAGAGATCACTGGAG 1049

Qy   476 accgtatctctgaagctcaacactgagcaattagaagaagtagtcctgtcaacaagaaggat 535
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1050 ACCGGGTCATGAAGTCTGAATACGGAAGATCACGAGGTCAGGAGACCTCTGAGCAAAAAGGGAT 1109

RESULT 4

ID   T18893   standard; cDNA; 4165 BP.
AC   T18893;
DT   05-JAN-1997 (first entry)
DE   Rat REK7 cDNA.
KW   REK7; eph-related tyrosine kinase receptor; AL-1; ligand;
KW   neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis;
KW   ss.
OS   Rattus sp.
FH   Key   Location/Qualifiers
FT   cds   541..3327
FT   /*tag= a
FT   signal_peptide 541..711
FT   /*tag= b
FT   mat_peptide   712..3324
FT   /*tag= c
PN   WO9613518-A1.
PD   09-MAY-1996.
PF   26-OCT-1995; U14016.
PR   27-OCT-1994; US-330128.
PR   07-JUN-1995; US-486449.
PA   (GETH) GENENTECH INC.
PI   Caras IW, Winslow JW;
DR   WPI; 96-239448/24.
DR   P-PSDB; R97853.
PT   AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
PT   treatment and diagnosis of neuronal disorders and
PT   angiogenesis-related conditions.
PS   Example 1; Page 47-49; 75pp; English.
CC   A cDNA clone (T18893) codes for rat REK7 (R97853), an eph-related
CC   tyrosine kinase receptor, for which AL-1 (see also W97854) is a
CC   ligand. It was isolated using degenerate receptor tyrosine kinase
CC   primers (T18894-96) to amplify cDNAs of an adult mouse hippocampal
CC   cDNA library. A PCR fragment was used as a probe to isolate the
CC   full-length REK7 cDNA from a rat hippocampal cDNA library. An
CC   REK-IgG fusion was used to screen cultured cell lines for surface
CC   expression of REK7-binding activity. Primers based on isolated
CC   ligands were used to amplify human breast carcinoma BT20 cell
CC   cDNA, and an amplified fragment was used to screen a human foetal
CC   brain cDNA library, leading to the isolation of AL-1 cDNA (T18897).
SQ   Sequence 4165 BP; 1157 A; 907 C; 1035 G; 1066 T;

Qy   536 tttatttggcattcaagatgttggttgcttgtcttggtgtcgtgtgagagtatact 595
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1110 TTTATCTTGCTTTCCAAGAATGTCGGTTGCTTTGTCCAAGAACGACCGGTACCCATGG 1169

Qy   596 tcaaaaagtgcccatttacagtgaagaaatctggctatgtgttccagaacatctggg 655
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1170 ATAAAAAATGTCCTTCTGTTAGTTAGACACATTTGGCTAGTTTTTCCCTGACACGATCACTGGAG 1229

Qy   656 actcccagtccc---tggtgggggtgagggggtttcttgttgtcaacaattctaaggaggaag 712
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 CAGAATTCATTCCCACGTTGCTAGAGGTGTCAGGCTCCTGGGTCAACCATTCTGTCGACAGACG 1289

Qy   713 atcctccaaggatgtactgcagtacagaaggcgaatggctcttgtaccccattggcaagtgtt 772
          | ||||||||||| |||||||||| |||||| || ||||||||||||||||||||||||||||
Db   1290 ATCCTCCAAAATGCATTCGACTGATCGTGCTGAAGGGAGGTGGTCTTGTTCCCATCGGGAAGTGCA 1349

Qy   773 cctgcaatgctggctatgaagaagagagagggtttttatgtgccaa 813
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1350 TGTGCAAGGCCGGGATATGAAGAGAGAAAAATGGTACCTGTCAA 1390

Query Match   43.28%; Score 351.4; DB 1; Length 4165;
Best Local Similarity   67.3%;
Matches 512; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy   56 tcggggaactgggattccgcagccttccaatgaagtcaatctactggattcaaaaacaattc 115
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   695 TCCGGACCCTTTGCCAGCCCCAGTAACGAAGTGAATTTATTGGATTCGCCCACTGTCC 754

```
RESULT 5

ID    V58192 standard; cDNA; 4322 BP.
AC    V58192;
DT    25-NOV-1998 (first entry)
DE    Mouse Bsk receptor-like tyrosine kinase cDNA clone.
KW    Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW    neurodegenerative disease; limbic system neuron regeneration;
KW    chromosomal abnormality; degenerative growth; development disorder;
KW    viral infection; bacterial infection; Alzheimer's disease; epilepsy;
OS    schizophrenia; stroke; cerebral ischaemia; ds.
OS    Mus sp.
FH    Key             Location/Qualifiers
FT    CDS             418..3051
FT                    /tag= a
FT                    /product= "Bsk"
FT                    /note= "receptor-like tyrosine kinase"
```

```
PA    (SCHU/) SCHULZ N T.
PA    (WOUD/) WOUDE G F V.
PA    (ZHOU/) ZHOU R.
PI    Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR    WPI; 98-541751/46.
DR    P-PSDB; W71628.
PN    US5814479-A.
PD    29-SEP-1998.
PF    11-JUN-1996; 673789.
PR    04-JAN-1994; US-177812.
PR    11-JUN-1996; US-673789.
PA    (KROM/) KROMER L F.
```

```
PT    Isolated nucleic acid sequence encoding protein - used in Bsk
PT    nucleic acid probes, used in detecting alterations in level of Bsk
PT    messenger-RNA in biological samples isolated from mammal afflicted
PT    with disease
PS    Claim 2; Fig 2; 72pp; English.
CC    The present sequence encodes mouse Bsk, which is a receptor-like
CC    tyrosine kinase. The nucleic acid sequence encoding Bsk can be used in
CC    Bsk nucleic acid probes, which can be used in detecting alterations in
CC    the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC    from a mammal afflicted with a disease, such as neurodegenerative
CC    diseases or disorders and neoplasms. The nucleic acid sequence can also
CC    be delivered into the limbic system of patients with limbic system
CC    neurodegenerative disease, disorder or injury, to promote or enhance
CC    limbic system neuron regeneration or growth. Such neurodegenerative
CC    diseases include, chromosomal abnormalities, degenerative growth and
CC    development disorders, viral infections, bacterial infections, brain
CC    injuries, neoplastic conditions, Alzheimer's disease, epilepsy,
CC    schizophrenia, or stroke and cerebral ischaemia.
SQ    Sequence 4322 BP; 1260 A; 887 C; 1055 G; 1120 T;
```

```
Query Match             43.0%;  Score 349.8;  DB 1;  Length 4322;
Best Local Similarity   67.1%;  Pred. No. 1.6e-105;
Matches  511; Conservative   0;  Mismatches  247;  Indels   3;  Gaps   1;
```

```
Db  1172  CAGAATTCATCACAGTTGTTGTTAGAGGTGTGTCAGGCTCCTGCTGCTCAACCAATTCGTGACAGATG  1231
          ||||||||||||| ||||| |||||||||||||||||||||||||||||||| |||||||||||||||
Qy   713  atcctccaaggatgtactgcagtacgaaggcgaatggctgctgtacccattggcaagtgtt  772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1232  ATCCTCCAAGATGCATTGCAGTTGCTGCTGAAGGGCTGGTTCCCATTGGGAAATGCA  1291
          ||||||||| ||||| ||||||||||||||||||||||||||||||||| ||||||
Qy   773  cctgcaatgctggctgtatgaagagaaaagagggtttatgtgccaa  813
          ||||||||||| ||||||||| ||||||||||| ||||||
Db  1292  TGTGCAAGGCTGGATATGAAGAGAGAAAAATGGTACCTACCTGCCAA  1332

RESULT  6
ID   V70208  standard; DNA; 3592 BP.
AC   V70208;
DT   11-FEB-1999  (first entry)
DE   Rat receptor tyrosine kinase Ehk-2 encoding DNA.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis; ss.
OS   Rattus sp.
FH   Key        Location/Qualifiers
FT   CDS        598..3444
FT              /*tag= a
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995;  469537.
PR   17-MAR-1995;  US-402247.
PR   26-JUL-1991;  US-735559.
PR   28-OCT-1993;  US-144992.
PR   06-JUN-1995;  US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Maslakowski P, Yancopoulos GD;
PT   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Claim 7; Fig 21; 194pp; English.
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence encodes rat Ehk-2.
SQ   Sequence 3592 BP; 946 A; 916 C; 894 G; 834 T;

Query Match       42.8%;  Score 348.2;  DB 1;  Length 3592;
Best Local Similarity     68.0%;  Pred. No. 4.9e-105;
Matches   501;  Conservative   0;  Mismatches 233;  Indels   3;  Gaps   1;

Qy    79  tccaatgaagtcaatctactggattcaaaaacaattcaagggggagctggggatctct  138
          |||||  |||||||||||||||||||| |||||||| |||||||||||||||||||||
Db   688  TCCAACCAAGTTGTGTTGCTTGATATACATTCACGTTGATGGGAGAACTAGGATGGGAAAACA  747

Qy   139  tatccatcacatgggtgggaagatcagtggtgttggatgaacattacacaccatcagg  198
          |||||  ||||||||||  |||| |||||||||||||| ||| |||||| ||||||||
Db   748  TATCCACTGAATGGGTGGGATGCCATTACTGAAATGGATGAACACACAGGCCCATACAT  807

Qy   199  acttaccaaggtgtgcaatgtcacagcacaccagtcaaaacaaattgtgctgagaacactgg  258
          |||||||||| ||||| ||||| ||||||||||| ||||||||||||||||| ||||||||
Db   808  ACATACCAGTGATGCAATGTCATCATGGAACTGGCTTCGTACTAACTGG  867

Qy   259  gtcccaaggaacctcaagctcaagaagattatgtgtggagctcaagttcactctacagagactgc  318
          |||||||||||| ||||||||||||| |||||||||| ||||||||||| ||||||||||||
Db   868  ATCCTCGTGATGCGTCCGAGAAAATCTATGTGTGGAAGTGAAGTCACATTGAGACGAAGTTGT  927

Qy   319  aatagacaattccaattggttttaggaactgtcaagaggagaacattcaacctgctactacatggag  378
          |||| ||||||||| |||| ||||||||| ||||||| ||||||||||||||||| |||||||
Db   928  AACAGCCATCCCATGGGGTCTTGGGAGTGTAAAGAACAATTACTCTGTATTATATTGAA  987

Qy   379  tctgatgatgatcatgggggtcttgggaactaaattcaaagattgacccatt  438
          |||| || | |||||| ||||||||||||| |||||| |||||||||||||||
Db   988  TCTGACGAATCCACGGAACTAAATTCAAGCCAATATATAAAGATTGACACAATT  1047
```

```
RESULT  7
ID   T02948  standard; cDNA; 3116 BP.
AC   T02948;
DT   16-APR-1996  (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK8 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK8;
KW   human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS   Homo sapiens.
FH   Key        Location/Qualifiers
FT   cds        34..2994
FT              /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995;  U04681.
PR   15-APR-1994;  US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-373799/48.
DR   P-PSDB; R85091.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 1; Page 57-62; 133pp; English.
CC   cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC   tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC   were isolated from a human foetal brain cDNA library using a directed
CC   PCR approach with primers (see T02960-61) based on conserved regions of
CC   receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC   extensive homology to the catalytic domain of chicken EPH-like
CC   receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC   EPH-like receptor. The isolated cDNAs are used for prodn. of
CC   recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC   to detect abnormalities in HEK receptor genes.
SQ   Sequence 3116 BP; 859 A; 720 C; 812 G; 725 T;

Qy   439  gcagctgatgaaagttcactcaaatgatcttgggggaccgtattctgaagctcaacact  498
          ||||||||||| |||||||||||||||||| |||||||||||||| |||||||||||||
Db  1048  GCTGCGGATGAGAGAGTTTTACTGGATTGGGTGGTCATCCGCATCCTTAAACTCAACACT  1107

Qy   499  gagattagagacaagtaggtcctgtcaacaagaagaaggatttttattggcaatttcaagatgtt  558
          |||||||||| ||||||||||| ||||||||||||||||| ||||||||||||||||||
Db  1108  GAAGTTCGTGAGGTGGGGCCAATAGAAAAGAAAAAAGGATTCATTTGGCTTTCAAGAATATT  1167

Qy   559  ggtgcttgttgttcttggtgtctgtgagagtatactccaaaaggtgcccatttacagtg  618
          ||| || || |||||||||||||| ||||| ||||| ||||||||||| |||||||||
Db  1168  GGAGCATGCATTGCTCGTGGTTGTGTCCAGTCGCGTGTGTTCTACAAAAAATGCCCCTTCACTG  1227

Qy   619  aagaaatctgtatgttcccaagacacggtaccca---tggactcccagtccctggttgggag  675
          ||||||||||||||||||||||||||||||||||   |||||||||||||||||| |||||
Db  1228  CGGAACTTGGCTATGTTTCCGATAACCATCCCAAGGGGTGACTCTTCTTCTTTGGTTGAA  1287

Qy   676  gttagagggttcttgttgtcaacaattctaaggaggaagatcctccaaggagtgtactgcagt  735
          ||||||||||||| ||||||||||||| |||||||| ||||| |||||||||| ||||||
Db  1288  GTGCGGGGCGCTCATGCCGTGAAGAGTTCTGAAGAGCGAGATACTCCTAAACTCGTTGGA  1347

Qy   736  acagaaggcgaatggcttgtaccagtgttcctgcaaatgctgtgatgaaga  795
          |||| |||| ||||||||| ||||||||||| ||||| |||||||| |||||
Db  1348  GCTGAGGAGAGGACTGGGCTCGTCCCTCTTGGAAGGTGTATCTGCACTACAGGGTATGAAGAGA  1407

Qy   796  agaggttttatgtgcca  812
          ||||||||||||||||||
Db  1408  ATCGAGAGGTTCTTGCCA  1424

Query Match       42.4%;  Score 344.4;  DB 1;  Length 3116;
Best Local Similarity     68.2%;  Pred. No. 8.3e-104;
Matches   509;  Conservative   0;  Mismatches 231;  Indels   6;  Gaps   2;

Qy    74  agcctccaatgaagtcaatctactggattcaaaaacaattcaagggggagctgggggctgga  133
          || | ||||||||||||||||||||||||||||||||||||||||||||| | | ||||||
```

```
Db  110 ACCCCGCGAATGAAAGTTACCTTATTGGAATTCCAGATCTCTGTTCAGGGGAGAACTTGGGTGGA 169
Qy
Db  134 tctc---ttatccatcacaatgggtgggaagagatcagtcagtgtggatgaacattacacac 190
            |||    |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy  170 TAGCAACCCTGGAAGGAGGGCTGGGAAGTGAGTGATATCAATGGATGAAAAAAATACAC 229
Db
Qy  191 ccatcaggacttaccaggttgtgcaagtcatggaccacacagtcaaaaacaatggctgagaa 250
Db  230 CAAATCCGAAACCTACCAAGTGTGCAATGTGATGAGGAACCCAGCCAGAGAATAAACTGGCTACGAA 289
Qy
Db  251 caaactgggtcccagcttcaggaagatttatgtgggagctcaacttcactctac 310
            |||| ||| ||| ||| | ||||||| || || ||||||||||| ||||| |||
Db  290 CTGATTGGATCACCCGAGAAGGGGCTCAGAGAGGGTGTATATTGGAGATTAAAATTCACCTTGA 349
Qy  311 gagactgcaatagacattccaatggttttattgaaacttgcaaggagacaattcaacctgtact 370
Db
Db  311 GGGACTCCAATAGTCTTCCGGCGGCGTCATGGGCGTCATGGGCGTCAAGGACGTTTAACCTGTACT 370
SQ
Qy  350 GGGACTCCAATAGTCTTCCGGCGGCGTCATGGGCGTCAAGGACGTTTAACCTGTACT 409
Db
Qy  371 acatggagtctgatgatgatcatgggggtgaaattctgagagacatcagttaacaaagattg 430
Db  410 ACTATGAATCAGACAACGACAAAAGACGTTTCATCAGAGAGAACCAGTTTGTTCATAAAAATTG 469
Qy  431 acaccattgcagctgatgaaagtttcactcaaatggatcttgggggacccgtattctgaagc 490
Db  470 ACACCATTGCTCCTGATGAAGAGCTTCACCCAAGTGGACAGAATCATGAAGC 529
Qy  491 tcaacactgaattagaagaaagttaggtcctgtcaacaagaaaggggatttcattttggcattc 550
Db  530 TGAACAACGAGATCCGGGAGATGTAGGGCCATTAAGCAAAAAGGGGTTTTACCTGCTTTTC 589
Qy  530 TGGTGGGAAGTTCGAGGGTCCGTCTGTCAACAACTTCAGAGAGATGTGCCAAAAAAGT 589
Qy  551 aagatgttggttgcttgtcttggtgtgtctgtgagagtatactttcaaaaagtgccaat 610
Db  590 AGGATGTTGGGGGCCTCGCATCCCCTGGTATCCGTGTGTTCTATAAAGTGTCCAC 649
Qy  611 ttacagtgaagaatctggctatgttttccagagaccagtccc---gactccagtccc 667
Db  650 TCACAGTCCGCAATCTGGCCCAGTTTCCTGACAATCACACAGGGCTTCGGATAGCGTCTCC 709
Qy  668 tggtgaaggttagaggggtcttgttgttgtctgtgaagtacttctttggcatttc 727
Db  710 TGGTGGGAAGTTCGAGGGTCCGTCTGTCAACAACTTCAGAGAGATGTGCCAAAAAAAGT 769
Qy  728 actgcagtacagaggcgaatggcttgtaccccatggcaagttgttcctgcaatgctggct 787
Db  770 ACTGTGGGCGCAGATGGTGAATGGCCTGCTGCTGTACTCCCATGGCCAACTGCCTGGGC 829
Qy  788 atgaagaaaggagggtttttatgtgccaa 813
Db  830 ATGAGGAGGACGGGACGGAGGAGAATGCCAA 855
```

RESULT   8

Query Match   42.4%;   Score 344.4;   DB 1;   Length 3348;
Best Local Similarity   68.2%;   Pred. No. 8.6e-104;
Matches 509;   Conservative   0;   Mismatches 231;   Indels   6;   Gaps   2;

```
Qy   74 agccttccaatgaagtcaatctactgggattccaaaaacaattcaaggggagactgggctgga 133
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 ACCCCGCGAATGAAAGTTACCTTATTGGAATTCCAGATCTCTGTTCAGGGGAGAACTTGGGTGGA 136
Qy  134 tctc---ttatccatcacaatgggtgggaagagatcagtcagtgtggatgaacattacacac 190
            |||    |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  137 TAGCAACCCTGGAAGGAGGGCTGGGAAGTGAGTGATATCAATGGATGAAAAAAATACAC 196
Qy  191 ccatcaggacttaccaggttgtgcaagtcatggaccacacagtcaaaaacaatggctgagaa 250
            |||||||||| |||||||||||| |||||||||||||| ||||||||||||||||||||| |||
Db  197 CAAATCCGAAACCTACCAAGTGTGCAATGTGATGAGGAACCCAGCCAGAGAATAAACTGGCTACGAA 256
Qy  251 caaactgggtcccagcttcaggaagatttatgtgggagctcaacttcactctac 310
            |||| ||| ||| ||| | ||||||| || || ||||||||||| ||||| |||
Db  257 CTGATTGGATCACCCGAGAAGGGGCTCAGAGAGGGTGTATATTGGAGATTAAAATTCACCTTGA 316
Qy  311 gagactgcaatagacattccaatggttttattgaaacttgcaaggagacaattcaacctgtact 370
            |||||  | ||||| || |||||||||||||||||||||||||| ||||| |||||||||||||
Db  317 GGGACTCCAATAGTCTTCCGGCGGCGTCATGGGCGTCAAGGACGTTTAACCTGTACT 376
Qy  371 acatggagtctgatgatgatcatgggggtgaaattctgagagacatcagttaacaaagattg 430
            ||||||| || ||| ||| ||||| | |||||||||||||||||| ||||||| ||||| |||
Db  377 ACTATGAATCAGACAACGACAAAAGACGTTTCATCAGAGAGAACCAGTTTGTTCATAAAAATTG 436
Qy  431 acaccattgcagctgatgaaagtttcactcaaatggatcttgggggacccgtattctgaagc 490
            |||||||||| |||||||| || |||||| |||||| || |||||||| ||| || |||||| |
Db  437 ACACCATTGCTCCTGATGAAGAGCTTCACCCAAGTGGACAGAATCATGAAGC 496
Qy  491 tcaacactgaattagaagaaagttaggtcctgtcaacaagaaaggggatttcattttggcattc 550
            |||||| ||| |||  |||  ||| |||| || ||||| |||||||| |||| ||||||||||||
Db  497 TGAACAACGAGATCCGGGAGATGTAGGGCCATTAAGCAAAAAGGGGTTTTACCTGCTTTTC 556
Qy  551 aagatgttggttgcttgtcttggtgtgtctgtgagagtatactttcaaaaagtgccaat 610
            ||||||||||  |||| || |||||||||| || ||||| || ||||||| || ||||||
Db  557 AGGATGTTGGGGGCCTCGCATCCCCTGGTATCCGTGTGTTCTATAAAGTGTCCAC 616
Qy  611 ttacagtgaagaatctggctatgttttccagagaccagtccc---gactccagtccc 667
            ||| ||| ||||||||||| |||||||| ||| | |||||||   |||||||||||||
Db  617 TCACAGTCCGCAATCTGGCCCAGTTTCCTGACAATCACACAGGGCTTCGGATAGCGTCTCC 676
Qy  668 tggtgaaggttagaggggtcttgttgttgtctgtgaagtacttctttggcatttc 727
            |||||  |||| |||||||| |||| ||||| ||||||||| ||||| ||||||||
Db  677 TGGTGGGAAGTTCGAGGGTCCGTCTGTCAACAACTTCAGAGAGATGTGCCAAAAAAGT 736
Qy  728 actgcagtacagaggcgaatggcttgtaccccatggcaagttgttcctgcaatgctggct 787
            ||||  ||||| || ||||||||| || ||||||||| |||||| ||| |||| |||||
Db  737 ACTGTGGGCGCAGATGGTGAATGGCCTGCTGCTGTACTCCCATGGCCAACTGCCTGGGC 796
Qy  788 atgaagaaaggagggtttttatgtgccaa 813
            ||||| || || ||| |||||||||||||
Db  797 ATGAGGAGGACGGGACGGAGGAGAATGCCAA 822
```

RESULT 9

```
ID    T02949    standard; cDNA; 4529 BP.
AC    T02949;
DT    16-APR-1996  (first entry)
DE    EPH-like receptor protein tyrosine kinase HEK11 cDNA.
KW    EPH-like receptor protein tyrosine kinase; PTK; HEK11;
OS    human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    cds             186..3182
FT                    /*tag= a
PN    WO9528484-A1.
PD    26-OCT-1995;
PF    14-APR-1995; U04681.
PR    15-APR-1994; US-229509.
PA    (AMGE-) AMGEN INC.
PI    Fox GM, Jing S, Welcher AA;
DR    WPI; 95-373799/48.
DR    P-PSDB; R85092.
PT    New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT    and related vectors, host cells, proteins, antibodies etc., used
PT    diagnostically and therapeutically to modulate receptor activation
PT    or prodn.
PS    Claim 1; Page 66-71; 133pp; English.
CC    cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC    tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC    were isolated from a human foetal brain cDNA library using a directed
CC    PCR approach with primers (see T02960-61) based on conserved regions of
CC    receptor PTKs and EPH-like receptor PTKs.  HEK5, HEK7 and HEK8 show
CC    extensive homology to the catalytic domain of chicken EPH-like
CC    receptors Cek5, Cek7 and Cek8.  HEK11 shows no homology to any known
CC    EPH-like receptor.  The isolated cDNAs are used for prodn. of
CC    recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC    to detect abnormalities in HEK receptor genes.
SQ    Sequence  4529 BP;  1149 A;  834 C;  1000 G;  1246 T;
```

```
Query Match  42.4%;  Score 344.4;  DB 1;  Length 4529;
Best Local Similarity  68.1%;  Pred. No. 1e-103;
Matches  512;  Conservative  0;  Mismatches  231;  Indels  9;  Gaps  2;

Qy   71  cgcagccttccaatgtcaactactggattcaaaaaaacaattcaagggagcctgggct  130
             ||||| |||| ||| |||  |||||||| ||||  ||| |  | |||||||  |
Db  265  CGCAGGCTGCAAGGAAGTACTACTGCTGGATTCTAAAGCACAACACAGAGTTGGAGT  324

Qy  131  ggatctcttatccacacatgggtcggggaagagaagatcaagtgtgtgtgatgaacactacacac  190
             | |  | |  |||||| |||||| || |||||   |||| || || ||||||  |||| |||||
Db  325  GGATTTCCTCTCCACCCAATGGGTCGGGAAGAAATTACTGGTTTGGAGGAACTATACC  384

Qy  191  ccatcaggaacttaccaggtgtgcaatgtcatgaaccacagtcaaacaacaattggctgagaa  250
             |||  | | || | |||||  ||  |||| ||||| |  || ||| ||||| ||||||||
Db  385  CGATACGAACATACCAGGTGTGCCAAGTCATGAGACCCACAAAACAACTGGCTCGGGA  444

Qy  251  caaactgggtccccaggaactcagctcagagaagcattattgtggagtcaactctac  310
             |||||| ||  ||| || |||| || ||||  |  | || | |||| ||| |||||
Db  445  CTAACTAGGATTTCAAAGGCAATGCACAAGGGATTTTTGTAGGAATTGAAAATTCACCCTGA  504

Qy  311  gagactgcaatagcatttcattgttttaggaactgcaggagacattcaaacctgtact  370
             ||| ||  || |   || | ||| ||||  |||||| | |  | |||| ||  ||| |
Db  505  GGGATTGTAACAGTCTTCCTGGAGTACTGGGAACTTCCAAGCAACATTTAATTGTGTACT  564

Qy  371  acatggagtctgatcaatggggtgaaattcgagagcatcagttcaagagatttg  430
             |||| | |  ||| | | |||||  ||| |||||||| | ||||||||  ||||
Db  565  ATTATGAAAACAGACTATGACACTGGGAGAAATATAAGAGAAAAACCTCTATGTAAAAAATAG  624

Qy  431  acacaccattgcagctgatgaaagtttcactcaaatgggacctgttctgaagc  490
             ||||| |||  | | | || ||||  | |  |   |||| |||| || |||| |
Db  625  ACACCAATTGCTGCACGATGAAAAGTTTTACCCAAGGGTGAACCTTGGTGAAAGAAGAAGATGAAGC  684
```

RESULT 10

```
ID    T02947    standard; cDNA; 3162 BP.
AC    T02947;
DT    16-APR-1996  (first entry)
DE    EPH-like receptor protein tyrosine kinase HEK7 cDNA.
KW    EPH-like receptor protein tyrosine kinase; PTK; HEK7;
OS    human eph-like kinase; therapy; diagnosis; vector; antibody; ss.
OS    Homo sapiens.
FH    Key             Location/Qualifiers
FT    cds             1..2976
FT                    /*tag= a
PN    WO9528484-A1.
PD    26-OCT-1995;
PF    14-APR-1995; U04681.
PR    15-APR-1994; US-229509.
PA    (AMGE-) AMGEN INC.
PI    Fox GM, Jing S, Welcher AA;
DR    WPI; 95-373799/48.
DR    P-PSDB; R85090.
PT    New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT    and related vectors, host cells, proteins, antibodies etc., used
PT    diagnostically and therapeutically to modulate receptor activation
PT    or prodn.
PS    Claim 1; Page 49-54; 133pp; English.
CC    cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC    tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC    were isolated from a human foetal brain cDNA library using a directed
CC    PCR approach with primers (see T02960-61) based on conserved regions of
CC    receptor PTKs and EPH-like receptor PTKs.  HEK5, HEK7 and HEK8 show
CC    extensive homology to the catalytic domain of chicken EPH-like
CC    receptors Cek5, Cek7 and Cek8.  HEK11 shows no homology to any known
CC    EPH-like receptor.  The isolated cDNAs are used for prodn. of
CC    recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC    to detect abnormalities in HEK receptor genes.
SQ    Sequence  3162 BP;  921 A;  667 C;  775 G;  799 T;
```

```
Query Match  42.2%;  Score 343.4;  DB 1;  Length 3162;
Best Local Similarity  66.6%;  Pred. No. 1.8e-103;
Matches  507;  Conservative  0;  Mismatches  251;  Indels  3;  Gaps  1;

Qy  551  aagatgttggttgcttgtcttgttgtctgtctgtgagagtatacttcaaaaagtgccat  610
             || || | | | | |||||| | ||| | | ||| |  |  ||  |   || |
Db  685  TTAACACTGAGGTGAGAGGAGAGGATTGGGACCTTTTGTCCAAAAAAGGGGATTCTATCTTGCCTTTC  744

Qy  611  ttacagtgaagaatctggctatgtttccagacacggtacc---catggacctcccagtccc  667
             | | |   |||  || || ||| |||| | ||| || |||       || ||  |||| ||
Db  745  AGGATGTAGGGGCTTGCATAAGCTTTGGTTTCGTCTCAAAGTGTACTACAAGAAGTGCTTGT  804

Qy  668  tgtgtggaggttgaggtgttcttgtctgtcaaccaatcttaagga------ggaaagatcctcaa  721
             | |||||||||   | ||| |||| | | || ||  ||  |||||      || ||| || |||
Db  805  CCATTATTGAGAACTTTAGCTACATCTTTCCAGATAACAGTGAACTCAGTGTCAGAGAATTTCCTCTT  864

Qy  722  ggatgtactgcagtacagaagcgaatggcttgtgtacccatggcaagtgttcctgcaatg  781
             ||| | || ||| ||||| |  | ||  | |||| |   |||||  |||| || ||||
Db  865  TAGTCGAGGTTGCAGGGACATGTGTCAGCAGTGCCAGAGGAAGAGCGGAAAACGCCCCCA  924

Qy  782  ctggcctatgaagaaaaggttttatgtgccaa  813
             ||| |  ||| ||| |||  || |  |  ||
Db  925  GGATGCACTGCACTGCAGAAGGAAGGTTAGTCGCCCATTGGAAAAARGTAGTCTGCAAAG  984

Qy        |  |||| ||  |||   | |||  || |||
Db  985  CAGGCTACCAGCAAAAAAGGAGACACTTGTGAA  1016
```

Qy 176 atgaacattacacaccatcaggacttaccaggtgttgcaatgtcatggaccagtcaaa 235
       |||| |||| | |||||| | | ||| ||| | | | | ||||| |||
Db 197 ATGAAAAATATGCCCCCTTATCCACACATACCAAGTAGTCCAAAGTGATGGAAAAATCAGA 256

Qy 236 acaaatggctgaagaacaaactcgggtccccaggaactcagctcagaagagattatgtggagc 295
       |||| ||||| | || | || || || | |||| | ||| || |
Db 257 ATAAACTGGCTTTTGACCAGTTGGATCTCCAATGAAGGTGCTTCCAGAATCTTCATAGAAC 316

Qy 296 tcaagttcacttcacgagactgcaatagcattccattggtttttaggaacttgcaaggaga 355
       |||| || || || || | | ||| | | ||| ||| || |
Db 317 TCAAATTTACCCTGCGGGGACTGCAACAGACGCTTCCTGGAGGACTGTAAGGAAA 376

Qy 356 cattcaacctgtactacatggtggagtctgatgtcatggggtgaaattcgagaggcatc 415
       || | | | || | ||| || | | ||| ||| | |
Db 377 CCTTTAATAATATGTTATTACTTGAGTCAGATGATCAGAAAACATCAAGGAGAAAAC 436

Qy 416 agtttacaaagattgacacattgcagctgatgaagtttcactcaaatggatcttgggg 475
       || ||| | || | ||| ||| || || ||||| ||| |
Db 437 AATACAACAAAATTGATACCAATTGCTCCCGATGAAAAGCTTTACAGAACTTGATCTTGGTG 496

Qy 476 accgtattctgaagcccaacacactgagaattagagaagtaggtccctgtcaacaagaaaggat 535
       || | | |||| || | | | || |||| | | |||||
Db 497 ACCGTGTTATGAAAACTGAAATACAGAGGTCAGAATGGACCTCTAAGCAAAAAGGGAT 556

Qy 536 tttattggcattcaagatgttggttcttgcttggttgtctgtggagtatact 595
       || || || |||| || | | | ||| || ||| | | |
Db 557 TTTATCTTGCTTTCAAGAGTTGGGTCGTTGCAATTGCTTGGTTTCGTGTGGCTGTGTATACT 616

Qy 596 tcaaaaagtgcccattacagtgaagaaatctcggctatgttgttccagacacgtaccatgg 655
       || || || | | |||| || | ||| | |||
Db 617 ATAAAAAATGCCCTTCCTGTGTGTACGACACTGGCTGGTGCTGTCTTCCCGTGACACACTC 676

Qy 656 actccagtccc----tggtggagggttagagaggtctgtgtcaacaattctaaggaggaag 712
       || | | | || || || || | || | | ||| | |
Db 677 CTGATTGTTCCCAATTGCTCGAAAGTGTCGGGGCTCCTCGTCGTCAACCATTCTCTGTGACCGATG 736

Qy 713 atcctccaaggatgtactgcgatacagaaggcgaatggcttgtctgtaccccattggcaagtgtt 772
       || | | | || ||| | || ||| | | |||
Db 737 AACCTCCCAAAAATGCACTGCACGCCCGAAGGGAGTGGCTGGTCCCATCGGGAAATGCA 796

Qy 773 cctgcaatgtctggctgtgatgaagaaagagggtttttatgtgccaa 813
       | ||| | || ||||| | | || |||
Db 797 TGTGCAAGGCAGGATATGAAGAGAGAAAAATGGCCACCTGTCAA 837

RESULT 11

ID T32962 standard; cDNA; 2323 BP.
AC T32962;
DT 11-NOV-1996 (first entry)
DE Mouse development kinase 1; MDK1 T2 clone.
KW Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW RTK; signal transduction; probe; diagnosis; gene therapy;
KW neurodegeneration; neuroproliferation; cancer; ss.
OS Mus sp.
FH Key   Location/Qualifiers
FT cds   233..2113
FT   /*tag= a
FT signal_peptide 233..316
FT   /*tag= b
FT mat_peptide   317..2110
FT   /*tag= c
FT misc_feature   2017..2113
FT   /*tag= d
FT   /note= "divergent sequence due to alternative
FT   splicing"

PI Clossek T, Millauer B, Ullrich A;
DR WPI; 96-333988/33.
DR P-PSDB; W03423.
PT New mouse developing kinase 1 gene - used for developing prods. for
PT diagnosis and treatment of abnormalities in signal transduction
PT pathways
PS Example 1; Page 115-116; 128pp; English.
CC cDNA cloning using adult mouse brains and Northern blotting
CC identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC coded for truncated versions (W03422 and W03423, respectively) of
CC the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC new member of the eck/eph family of receptor tyrosine kinases.
CC MDK T1 and T2 each possess the entire ectodomain, the transmembrane
CC domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC but lack the catalytic tyrosine kinase domain. They may have a
CC modulatory function. The cDNA clones can be used to produce MDK1
CC T1 and T2, which are useful for screening potential agents for
CC treatment of diseases characterised by abnormal signal transduction.
SQ Sequence 2323 BP; 680 A; 497 C; 609 G; 537 T;

Query Match   41.2%; Score 334.8; DB 1; Length 2323;
Best Local Similarity 67.3%;
Matches 506; Conservative 0; Mismatches 237; Indels 9; Gaps 2;
                  Pred. No.1.1e-100;

Qy 71 cgcagccttccaatgaagtcaatcctactggtgattcaaaaacaattcaagggggactgggct 130
     || | | || || | || || | | |||| |
Db 312 CGCAGGCGTCCGAAGGAGAACTACATTATTACTGGACTGCGAAAGCACACAACAGAATTGGAAT 371

Qy 131 ggatctcttatccatctcacatgggtggagaagatcagtggtgtggatgaacattacac 190
     |||| | | | | || | ||| | | ||
Db 372 GGATTCTCTCTCCACCCAGTCGGTGGGGAGGAAAATTAGTGGTTGGATGAGAACTACACTC 431

Qy 191 ccatcaggagacttaccaggtgttgcaatgtcatggaccagtcaaaaaccaattggctgaaa 250
     || ||| || ||| | | | | || |||| |
Db 432 CGATAAGAACANTACCAGGTGTGCCCAGGTTCATGGAGCCCAACCAAACTGGCTGCCGGA 491

Qy 251 caaactgggtcccaggaactcagctcagaagagattttatgtggagctcaagtcacctctac 310
     || ||| || | || | | | | ||| |
Db 492 CTAACTGGATTTCTAAAGGCACAAGCACAAAAGGATTTTTGTGAATGAAATTCACCTTGA 551

Qy 311 gagactgcaatagcatccattggtttttaggaacttgcaaggagacattcaacctgtact 370
     || || || | || | || | | | | |
Db 552 GGGATTGTAATAGTCTTCCCGGAGTCCTGGAGACTTGCAAGGAAACGTTTAATTGTGTACT 611

Qy 371 acatggtggagtctgatgatgatgtgggtgaaatttcgagagcatcagtttacaaagattg 430
     | ||| | || || | || | || | |
Db 612 ATTATGAAACAGACTACGACACCGGCAGGAATATACGAGAAAAACCTTTAIGTTAAAATAG 671

Qy 431 acacacattgcagctgatgaaagtttcactcaaatggatctcttgggggtccgtatcctgaagc 490
     || | || | || ||| | | | |||
Db 732 ACACCAATTGCTCGAGATGAAAGTTTCACACAAGGTGACCTTGGTTGAAAGAAGATGAGAGC 731

Qy 491 tcaacactgagaattagagaagtaggtcctgtcaacaagaaaggggatattcaacctgtact 550
     || || | | | | || || | || |
Db 732 TGAACACTGAGGTGAGAGAGATTGGACCTTGTGCCAAAAAGGGATTCTATCTTGCCTTTC 791

Qy 551 aagatgttgggtgcttgttgttgccttggttgtctgtgagtatactacttcaaaaagtgccat 610
     || | || | || | | || | | |
Db 792 AGGATGTAGGGGGCTTGCATAGCACATTGGTTTCTCAAAGTGTACTACAAGAGTGCTGGA 851

Qy 611 ttacaagtgaagaatctggctatgttccagacacgtacc---catggacctcccagtgccc 667
     || | | || || || | | || |
Db 852 CCAITGTTGAGAACTTAGCTGTGTCTTTCCAGATACAGTGACTGGTTCGGAATTTTTCCCT 911

Qy 668 tggtggagggttagagaggtctgtgtcaacaattctaagga------ggaagatcctcaa 721
     || || || || | | | | || |||
Db 912 TAGTCGAGGTCCGTGGGACACATTGGTCACCAGTGCCGAGGAAGAGGCAGAAAAAATTCCCCA 971

Qy 722 ggatgtactgcgatacagaaggcgaatggcttgtctgtaccccagtgccaatg 781
     || | | ||| | | || | || | | |
Db 972 GAAAGCAATTGCAGTGCAGCAGGAGAGTGGCTAGTACCCATTGGAAAAATGCATCTGCAAAG 1031

RESULT 12

```
Qy   782 ctggctatgaagaagaggtttatgtgccaa 813
         ||||| ||  ||| ||  |||  || ||||
Db  1032 CAGGCTATCAGCAAAAAGGGGACACTTGCGAA 1063
```

```
RESULT 12
ID   T32961
AC   T32961;
DT   11-NOV-1996  (first entry)
DE   Mouse developmental kinase 1 MDK1 T1 clone.
KW   Mouse developmental kinase 1; MDK1; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   cds             233..2065
FT                   /*tag= b
FT   signal_peptide  233..316
FT                   /*tag= a
FT   mat_peptide     317..2062
FT                   /*tag= c
FT   misc_feature    2031..2065
FT                   /*tag= d
FT   polya_signal    2839..2843
FT                   /*tag= e
FT                   /note= "divergent sequence due to alternative
FT                   splicing"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; US-368776.
PR   03-JAN-1996; U00419.
PA   (PLAC-) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 111-12; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC   coded for truncated versions (W03422 and W03423, respectively) of
CC   the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC   new member of the eck/eph family of receptor tyrosine kinases.
CC   MDK1 T1 and T2 each possess the entire ectodomain, the transmembrane
CC   domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC   but lack the catalytic tyrosine kinase domain. They may have a
CC   modulatory function. The cDNA clones can be used to produce MDK1
CC   T1 and T2, which are useful for screening potential agents for
CC   treatment of diseases characterised by abnormal signal transduction.
SQ   Sequence 2901 BP;
DR   WPI; 96-333988/33.
DR   P-PSDB; W03422.
```

```
Query Match    41.2%;  Score 334.8;  DB 1;  Length 2901;
Best Local Similarity  67.3%;
Matches 506;  Conservative 0;  Mismatches 237;  Indels 9;  Gaps 2;
```

```
Qy    71 cgcagcgcttccaatgaagtcaatctactggattcaagaaaacaattcaagggagctgggct 130
Db   312 CGCAGGCTCGGAAGGAGAAGTACTATTACTTGGACTCGAAAGCACAACAAACAGAAATTGGAAT 371

Qy   131 ggatcctcttatccatcacacatgggtggggaaagaaaagatacaagttgttgttgatgaacattacacac 190
Db   372 GGATTTCCTCTCTCCACCCAGTCGGTGGGGAAGAAATTAGTGGTTTGGATGAGAACTACACAC 431

Qy   191 ccatcacaggaactaccaggttgtgccaatgtggaccacacgtcaaaaacaattggctgagaa 250
Db   432 CGATAAGAACATACCAGGTGTGCCAGGTCATGGAGGCCCACACAGAACACAACTGGCGCTGCGGA 491

Qy   251 caaactgggtccccaggaactcagctccagaagaagatttatgtgggagctcaagttcacttctac 310
```

RESULT 13

```
Db   492 CTAACTGGATTCTAAAGGCANCACGCACAAAAGGACGATTTTTGTAGAAATTGAAATTCACCTTGA 551

Qy   311 gagactgcaatagcattccaattggtttgacttgcaaggaggagacattcaacctgtact 370
Db   552 GGGATTGTAAATAGTCTTCCCGGAGTCCTGGGGAACTTGCAAGGAAAACCTTTATTGTTACT 611

Qy   371 acatggagtctgatgatcatgtgggtgaaattgaagatctggggaccatcagttacaaagattg 430
Db   612 ATTATGAAACAGACTACCACACACCGGCACAGGAAATATACCGAGAAAACCTTTATGTGTAAAATAG 671

Qy   431 acaccattgcagctgatgaaagtttcactcaaatggatcctggggaccgtatctgaagc 490
Db   672 ACACCATTGCTCCGAGATGAAAGTTTCACACAACAGTGAACAGTGAACTTGGTGAAAGAGATGAAGC 731

Qy   491 tcaacactgagattagagagagtaggtcctgtcaacacaagaaggggatttatttggcattc 550
Db   732 TGAACACTGAGGTGAGAGAGGATTGAGAGATTGCCAAAAGGGAATTCTTATCTTGCCGTTTC 791

Qy   551 aagatgtttggtcgcttgtgttgcctttgtgtctgagtattacttcaaaaagtgccat 610
Db   792 AGGATTGTAGGGGGCTTGCATTAGCATTGGTTTCGTCGTCAAAGTGTACTACAAGAGTGCTGGA 851

Qy   611 ttacagtgaagaatctggctatgttccagaacacggtacc---catggactcccagtccc 667
Db   852 CCATTGTTGAGAACTTAGCTGTCGTCTTTCCAGATACAGCAGTGATGACTCGGTTCGGAATTTCCTCCT 911

Qy   668 tggtgggggttagaggtgtcttgtcaacaatctaagga------ggaagatcctcaa 721
Db   912 TAGTCGAGGTCCGTGGGACATGTGTCAGCAGTGCCGAGGAAGGCAGAGAAAATTCCCCA 971

Qy   722 ggatgtactgcagtacaagaaggcgaatggcttgtacccatggcaagtgttcctgcaatg 781
Db   972 GAATGCAATTGCAGTGCACAAGAAGAGAGTGGCTAGTACCCATTGGAAAATGCACTCTGCAAAG 1031

Qy   782 ctggctatgaagaagaggtttatgtgccaa 813
Db  1032 CAGGCTATCAGCAAAAAGGGGACACTTGCGAA 1063
```

```
RESULT 13
ID   T32960
AC   T32960;
DT   11-NOV-1996  (first entry)
DE   Mouse developmental kinase 1 cDNA.
KW   Mouse developmental kinase 1; MDK1; receptor tyrosine kinase; RTK;
KW   signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   cds             233..3229
FT                   /*tag= b
FT   signal_peptide  233..316
FT                   /*tag= a
FT   mat_peptide     317..3226
FT                   /*tag= c
FT   3'utr           3230..3628
FT                   /*tag= e
FT   polya_signal    3568..3573
FT                   /*tag= e
FT   3'utr           3627..4304
FT                   /*tag= f
FT                   /note= "alternative 3'-untranslated region,
FT                   found in one isolated clone"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; US-368776.
PR   03-JAN-1996; U00419.
PA   (PLAC-) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
DR   WPI; 96-333988/33.
DR   P-PSDB; W03421.
```

```
PT  New mouse development kinase 1 gene - used for developing prods. for
PT  diagnosis and treatment of abnormalities in signal transduction
PS  pathways
PS  Example 1; Page 103-105; 128pp; English.
CC  A cDNA clone (T32960) codes for mouse developmental kinase 1 (MDK1)
CC  (W03421), a new member of the eck/eph family of receptor tyrosine
CC  kinases (RTKs). To isolate the clone, cDNA from mouse embryos was
CC  subjected to PCR amplification with primers based on conserved
CC  motifs (see also W03426-27) of RTKs. An amplified fragment was used
CC  to screen an 11.5-day-old mouse embryo and an adult mouse brain cDNA
CC  library to obtain the MDK1 clone. 4 Sequence variants (see also
CC  T32961-62 and W03422-25) of MDK1 were also identified. MDK1 nucleic
CC  acids can be used for the recombinant prodn. of MDK1, as probes to
CC  detect MDK1, and for the gene therapy of diseases involving
CC  abnormalities in signal transduction, such as neurodegenerative and
CC  neuroproliferative disorders and cancer.
SQ  Sequence 4304 BP; 1297 A; 875 C; 1043 G; 1089 T;

Query Match         41.2%;  Score 334.8;  DB 1;  Length 4304;
Best Local Similarity   67.3%;
Matches  506;  Conservative    0;  Mismatches 237;  Indels   9;  Gaps   2;

Qy   71 cgcaagccttcaatgaagtcaactctactggattcaaaaacaattcaagggaggctgggct 130
        |||||| || |||  ||||| ||||||||||| | |||||  ||||||||||  ||||| ||
Db  312 CGCAGGCTGCGAAGGAAGTACTATTACTGGACTCGAAAGCCACACAACAGAATTGGAAT    371

Qy  131 ggatctcttatccatccacatgggtgggaagaagatcagtgggtggtgatgaacattacac 190
        ||||| || | || |||| |||||||||||||| ||| ||| ||||||| |||||||||||
Db  372 GGATTCCTCTCCCACCCAGTGGGGTGGGAAGAAATTAGTTGGTTTGGATGAGAACTAACTC 431

Qy  191 ccatcaggagacttaccaggtgtgcaaatgtcatggaccacagtcaaaacaaattggctgagaa 250
        | ||| ||| ||| ||||||||||| |  ||||||| |||| |||||||| | ||||||||| |
Db  432 CGATAAGAACAATACCAGGTGTGCCAGGTCATGGAGCCCAAACAACTGGCTGCCGGA       491

Qy  251 caaactgggtccccaggagaactcagtcaagagagaagatttatgtgagctcaagttcaactctac 310
        |  |||||| || || ||| | ||| | ||| |   ||| || ||  |||||||||| || |||||
Db  492 CTAACTGGATTTCTTAAAGGCACGCACAAGGAATTTTGTAGAATTGAAATTCACCCTTGA    551

Qy  311 gagactgcaatagcattccaatccatggttcttaggaactttgcaagttgcaaggag 370
        |  ||||| || || ||| |  |||||||||||||||||| |||||| ||||||||
Db  552 GGGATTGTAAATAGTCTTCCCGGAGTCCTGGGAACTTCCAAGAAACGTTTAATTTGTACT 611

Qy  371 acatggggtctgatgggtgggtgaaatttcagtttacaaagatg 430
        | ||| |||||||||||| ||||| ||||| |||| ||||||||
Db  612 ATTATGAAAACAGACTACGACCACCGGCAAGAATATACGAGAAAAACCTTTAIGTTAAAAIAG 671

Qy  431 acaccaccattgcagctgtgaaagtttcactcaaatggatcttgggggaccgtattctgaagc 490
        ||||| | |||||||||| | |||||||||| |||||||||||||| ||||| |||| ||||
Db  672 ACACCAATTGCTGCGAGATGAAAAGTTTCACACAAGGTGACCTTGGTGACAAGANTGAAGC  731

Qy  491 tcaacacactgagttagagaagtaggtcctgtgtcaacaagaaggagtataca 550
        |||| || ||||| ||||||| ||| | | |||||||||||| || |||| |
Db  732 TGAACACTGAGGTGAGAGAGAATTGGGACCTTGTTCGTCCAAAAAAGGGATTCTATCTTGCCTTTC 791

Qy  551 aagatgttggttgcttgtctgtctgtgagagtatacttcaaaagtgccaat 610
        |||||| | |   ||| ||   ||| ||||||||||| |||||||| || |
Db  792 AGGATGTRAGGGGGCTTGCAATAGCATTGGTTTCTCGTCAAAAGTGTACTACAGAAGTGCTGGA 851

Qy  611 ttacagtgaagaatctgctatgtttcagaacacgtacc---catggactcccagagtcc 667
        ||  ||||||||| | || |||||||||||  || | |   ||||||||  |||  ||
Db  852 CCAITGTTGAGAGACTTAGCTGCTGTTTCCAGATACAGTCGGTCGGAATTTTCCTCT       911

Qy  668 tggtggagttagaggggtccttgtgtcaacaattctaagga------ggaagatcctccaa 721
        | || ||| |||| ||| |  | |||||||| |  ||| ||      ||||||| ||| ||
Db  912 TAGTCGAGGTCGCGTGGGACATGTGTCAGCAGTCGCGAGGAAGCAGAAAAAATTCCCCA     971

Qy  722 ggatgtactgcagtacagaaggcggatggcttgtaccccaagtgccaagtgt 781
        |||||| |||||| | |||||| || ||||| ||||||||| |||| ||||
Db  972 GAATGCAITGCACAGTCGCGACATGGCTAGTACCCATTGAAAAATGCAITCTGCAAAG     1031

Qy  782 ctggctcatgaagaaagagggtttatgtgccaa 813
        | ||||| |||||| | ||||| |||| |||||
Db 1032 CAGGCTATCAGCAGAAAAAGGGACACTTGCGAA 1063
```

```
RESULT 14
ID  Q90662 standard; cDNA; 3056 BP.
AC  Q90662;
DT  11-NOV-1995 (first entry)
DE  Eph-related PTK Cek7/ cDNA.
KW  Cek7/; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW  prognosis; ss.
OS  Gallus sp.
OS  Gallus sp.
FH  Key             Location/Qualifiers
FT  cds             2..2134
FT                  /tag= a
PN  WO9915375-A.
PR  08-JUN-1995;
PF  07-SEP-1994; U10140.
PF  03-DEC-1993; US-162809.
PR  (LJOL-) LA JOLLA CANCER RES FOUND.
PA  Pasquale EB, Sajjadi FG;
PI  WPI: 95-215256/28.
DR  P-PSDB: R75714.
PT  Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT  cancer.
PS  Claim 2; Page 106-109; 129pp; English.
CC  Novel EPH-related PTK cDNA clone Cek7 (given in Q90653) and its
CC  variant Cek7+ (Q90661) were isolated from a chick embryo library in
CC  lambda gt11, and another variant, Cek7 (Q90662), from a chick
CC  embryonic brain cDNA library in lambda gt11. The variants may originate
CC  via alternative splicing of the same gene. Cek7 had the lowest level
CC  of expression of 7 novel Eph-related kinases examined and was barely
CC  detectable in adult tissues.
SQ  Sequence 3056 BP; 871 A; 661 C; 762 G; 762 T;

Query Match         30.2%;  Score 245.2;  DB 1;  Length 3056;
Best Local Similarity   68.2%;  Pred. No. 5.5e-71;
Matches  356;  Conservative    0;  Mismatches 163;  Indels   3;  Gaps   1;

Qy  295 ctcaagttcactctacgagactgcaatagcattccatggttcttaggaacttgcaagggg 354
        |||||  ||||| |||  ||||||  |  || |||||  | |||| |||||| |||||| |
Db    2 CTCAAATTCACCCTGAGGGACTGTRAACAGCCTTCCAGGAGGACTTGGGACTTGCAAGGAG 61

Qy  355 acattcaacctgtactacatgagagtctgatgatgatcatgggtgggtgaaatttcgaggcat 414
        ||| | ||| |||||||| |||||| | | |||||||  |||||| ||  || | ||| ||
Db   62 ACTTTTAACATGTACTACTTTGAGTCAGAATGATGAAGGATGGAGGAACATCAGAGAGAAT 121

Qy  415 cagttacaaagattgacaccattgcagctgatgaaagttcacttcaaatggatcttggg 474
        |||| ||||||||| || ||||||| |  |||||||| |||| |||| |||||| |||
Db  122 CAGTACATCAAAGATARGATAACCATTGCCTGCTCGATGAGAGCTCGGC           181

Qy  475 gaccgtatctgaagctcaacactgagattagagaagaaggatcctgtgtcaacaagaaggga 534
        ||| | ||| |||| ||| || |||| ||||| |||||| | ||| ||||||||| ||||
Db  182 GACAGAGTTAMGAAGTTAAACACAGAAGGTGAGAGAGATAGTGGAGCTTCGGC       241

Qy  535 ttttattttggcattcaagatgttggtcttgtgtctgtgagagtatac 594
        |||| |  ||| ||||| ||||| || ||| ||| | |||||||||||
Db  242 TTTTACCTTGCTTTCCAGGATGTGGGGCCTGCAATGCCCTGGTCTCTGTGCGTGTGTAC 301

Qy  595 ttcaaaaagtgccacattacagtgaagaatctggctatgtttcagaacacggtaccatg 654
        | || |||||||| || || ||||||||| | |||| |||||||||  ||  | || ||
Db  302 TACAAGAAAATGCCCATCAGTGAATCGGCCACGCTTTCCAGAATACCATCACACAGGA   361

Qy  655 ---gactcccagtccctgtggaagttagaggggtcttgtgtcaacaattctaaggaggaa 711
           ||||||||| |||||||| |||| |||| ||| | | |||||||| |  |||||||
Db  362 GCAGANTTCCTCCGCACCTGCTAGAAGTGTCAGGCGTCAACAAGGCTGACTGAT       421

Qy  712 gatcctccaaggatgttactgcagtacagaaggcgggaatggccttgtaccccaagtgt 771
        ||||| ||||||||| |||||||| | |||||| ||  ||||| ||||||||| |||
Db  422 GAGGCACCAAAGATGCACACTGCAGTTCAGAGGGAGAAATGGCTGGTGCCCAGTGCCAGAGTGT 481

Qy  772 tcctgcaatgctgctatgaagaagagggtttatgtgccaa 813
        || || |||||||| |||||||| ||||| |||| |||||
Db      ...                                    813
```

```
Db    482 TTGTGCAAGGCAGGGTACGAGGAGAAGAACAACACCTGCCAA 523

RESULT  15
Q90653
ID   Q90653 standard; cDNA; 3059 BP.
AC   Q90653;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek7 cDNA.
KW   Cek7; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis; ss.
OS   Gallus sp.
FH   Key            Location/Qualifiers
FT   cds            2..2170
FT                  /*tag= a
PN   WO9515375-A.
PD   08-JUN-1995.
PF   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;
DR   WPI; 95-215256/28.
DR   P-PSDB; R75705.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Claim 2; Page 44-47; 129pp; English.
CC   Novel EPH-related PTK cDNA clone Cek7 (given in Q90653) and its
CC   variant Cek7+ (Q90661) were isolated from a chick embryo library in
CC   lambda gt11, and another variant, Cek7' (Q90662), from a chick
CC   embryonic brain cDNA library in lambda gt11. The variants may originate
CC   via alternative splicing of the same gene. Cek7 had the lowest level
CC   of expression of 7 novel Eph-related kinases examined and was barely
CC   detectable in adult tissues.
SQ   Sequence 3059 BP;     860 A;     632 C;     758 G;     809 T;

Query Match     30.2%;
Best Local Similarity   68.2%;
Matches   356;  Conservative   0;  Mismatches  163;  Indels   3;  Gaps   1;
                                   Score 245.2;  DB 1;  Length 3059;  Pred. No. 5.5e-71;

Qy    295 ctcaagttcactctacgagactgcaatagcattccattggtgtttaggaacttgcaaggag 354
          |||| |||| || ||  | ||| | ||||||| ||||| ||||| ||||| |||||||||||
Db      2 CTCAAANTTCACCCTGAGGGACTGTAACAGCCTTCCACGAGGACTTGGGACTTGCAAGGAG 61

Qy    355 acattcaaacctgtactacatggagtctgatgatgatcatgggtgaaattccgaggcat 414
          || ||| ||| ||| |||| ||||||| ||||||||  | ||||  |||| ||| ||||||
Db     62 ACTTTTAAACATGTACTACTTTGAGTCAGATGATGAAGATGGGAGGAACATCAGAGAGAAT 121

Qy    415 cagtttacaaagttgacacacattgcagcctgatgaaagtttcactcaaatggatcttggg 474
          |||| ||| || || |  || ||||| ||  | |||| |||||||| || |||||| |||
Db    122 CAGTACATCAAGATAGATAACATTGCTGCTGAATGAGAGCTTCACGGAGTTGGACCTCGGC 181

Qy    475 gaccgtatctgaagctcaacactgagattagaaggtaggtccgtgtcaacaagaaggga 534
          ||| | | | ||  || |||||| || |||  |||||||| || |  |||| ||| ||||
Db    182 GACAGAGTTATGAAGTTAAACACAGAAGATGAGAGANTGTGGCCCTCAACAAAAAAAGGA 241

Qy    535 ttttatttggcattccaagatgttggtggctgttgtgttctggtctgttgaagagtatac 594
          ||||   |||| |||||| ||||| || | ||  || | |  ||||| |||| ||| |||
Db    242 TTTTAACTTGCTTTCCAGGATGTGGGCGCCCTTGCATTGCCCTGGTCTCTGTGGCTGTGTAC 301

Qy    595 ttcaaaaagtgcccattacagtgaagaatctggctatgtgttccagacacggtaccatg 654
          | ||| ||| ||||| | || ||  ||  |  ||| | || |||||| |||| |||||
Db    302 TACAAGAAAATCCCATCCAGTGATCCCGCAAACCTGGCACGCTTTCCACGATACCATCACAGGA 361

Qy    655 --gactcccagtcctggtggagttagaggtcttgtgtcaacaattctaaggaggaa 711
            || | ||| | |||| || ||| |  | ||  ||||||| |  || || ||||
Db    362 GCAGAATTCCTGCAGCTGCTAGAAGTGTCAGCGGTCGTGTGTCAACCACTCACTGACTGAT 421

Qy    712 gatccccaaggatgtactgcagtacaagaggcgaatgcggcttgtgtaccaccaagtgt 771
          ||  | ||||||||||||||||| ||| | || | ||||  |||| |  |||| ||||| 
Db    422 GAGGCACCAAAGATGCACTGCAGTTCAGCATTGGGTGCCCCCATTGGGAAGTGT 481

Qy    772 tcctgcaatgctggcctattgaaagaaagaggtttatgtgccaa 813
          | |||||| | ||||| || |||| ||||| |||||||||||||
Db    482 TTGTGCAAGGCAGGGTACGAGGAGAAGAACAACACCTGCCAA 523
```

Search completed: May 15, 2000, 12:08:19
Job time: 4578 sec

This Page Blank (usproj

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:　　　May 15, 2000, 06:32:53 ; Search time 1225.64 seconds
　　　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　　　　-645.279 Million cell updates/sec

Title:　　　　US-09-104-340-5
Perfect score:　813
Sequence:　1 atggattgtcagctctccat.........aaagaggtttatgtgccaa 813

Scoring table:　IDENTITY_NUC
　　　　　　　　Gapop 10.0 , Gapext 1.0

Searched:　　882769 seqs, -48639572 9 residues

Total number of hits satisfying chosen parameters:　1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
　　　　　　　　Listing first 45 summaries

Database :　GenEmbl:*

1: gb_bal:*	45: gb_htg7:*
2: gb_ba2:*	46: em_htg1:*
3: gb_om:*	47: em_htg2:*
4: gb_ov:*	48: em_htg3:*
5: gb_pat:*	49: em_hum5:*
6: gb_ph:*	50: gb_pl3:*
7: gb_pl1:*	51: gb_pr5:*
8: gb_pl2:*	52: gb_htg8:*
9: gb_pr1:*	53: gb_htg9:*
10: gb_pr2:*	54: gb_htg10:*
11: gb_pr3:*	55: gb_htg11:*
12: gb_ro:*	56: gb_htg12:*
13: gb_sts:*	57: gb_htg13:*
14: gb_sy:*	58: gb_htg14:*
15: gb_un:*	
16: gb_vi:*	
17: em_fun:*	
18: em_hum1:*	
19: em_hum2:*	
20: em_in:*	
21: em_om:*	
22: em_or:*	
23: em_ov:*	
24: em_pat:*	
25: em_ph:*	
26: em_pl:*	
27: em_ro:*	
28: em_sts:*	
29: em_sy:*	
30: em_un:*	
31: em_vi:*	
32: gb_htg1:*	
33: gb_htg2:*	
34: gb_in1:*	
35: gb_in2:*	
36: em_ba1:*	
37: em_ba2:*	
38: em_hum3:*	
39: em_hum4:*	
40: gb_pr4:*	
41: gb_htg3:*	
42: gb_htg4:*	
43: gb_htg5:*	
44: gb_htg6:*	

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result NO.	Score	% Query Match	Length	DB	ID	Description
	1	813	100.0	3132	12	A28003	A28003 H.sapiens H
	2	813	100.0	3132	4	I68018	I68018 Sequence 9
	3	813	100.0	3149	9	HUMHEK	HUMHEK Human recep
	4	678.6	83.5	3077	12	RNU69278	U69278 Rattus norv
	5	661	81.3	2032	12	MUSMEK4SE	M68515 Mouse eph-r
c	6	661	81.3	3197	12	MUSMEK4	M68514 Mouse eph-r
	7	546.6	67.2	3241	4	CHKCEK4	M68513 Chicken eph
	8	545	67.0	3254	5	I15007	I15007 Sequence 15
	9	362.8	44.6	3600	4	CHKCEK8	D38174 Gallus gall
c	10	359.4	44.2	3943	12	MMU58332	U58332 Mus musculu
	11	351.4	43.2	3531	12	RNEHK1	X78689 R.norvegicu
	12	351.4	43.2	3906	5	AR062744	AR062744 Sequence
	13	351.4	43.2	4165	5	AR025488	AR025488 Sequence
	14	351.2	43.2	4242	12	MMSEK	X65138 M.musculus
	15	351.2	43.2	4242	12	S57168	S57168 Sek-Eph-rel
	16	349.8	43.0	4322	4	AR043381	AR043381 Sequence
	17	349.8	43.0	4322	12	MMU07357	U07357 Mus musculu
	18	348.8	42.8	3592	5	AR062743	AR062743 Sequence
	19	345	42.4	3162	10	HUMRPTKB	L36644 Homo sapien
	20	345	42.4	3903	10	HSEHK1	X95425 H.sapiens m
	21	344.4	42.4	3107	10	HUMRPTKC	L36645 Homo sapien
	22	344.4	42.4	3348	10	I44522	I44522 Sequence 34
	23	344.4	42.4	4523	10	HUMRPTK	L36642 Homo sapien
	24	341.4	42.0	3042	4	XLSEK1	X91191 X.laevis mR
	25	341.4	42.0	3193	4	XELPAGAAA	L26099 Xenopus lae
	26	334.8	41.2	2323	12	MMKINIT2	X79084 M.musculus
	27	334.8	41.2	2901	12	MMKINIT1	X79083 M.musculus
	28	334.8	41.2	4304	12	MMKIN1	X79082 M.musculus
	29	331.6	40.8	2006	12	RNU21955	U21955 Rattus norv
c	30	331.6	40.8	3208	12	RNU21954	U21954 Rattus norv
	31	331.6	40.8	3943	11	GGY14271	Y14271 Gallus gall
	32	330.6	40.7	77197	11	U90093	U90093 Human Chrom
	33	329	40.5	4124	4	GGCEK7B	AC021499 Gallus gall
	34	321.2	39.5	76022	45	AC021499	AC021499 Homo sapl
	35	316.2	38.9	4737	12	MMU72207	U72207 Mus musculu
c	36	307.4	37.8	132805	41	AC009425	AC009425 Homo sapl
	37	305	37.5	4577	4	HSJ189K14	AL121966 Homo sapl
	38	305	37.5	2640	4	DRU82295	U89295 Danio rerio
	39	297.2	36.6	160703	10	DRAJ5030	AJ005030 Danio rer
c	40	276.6	34.0	3056	10	HS61A9	AL035703 Human DNA
	41	245.2	30.2	3059	5	I15010	I15010 Sequence 21
	42	245.2	30.2	3125	5	I15001	I15001 Sequence 3
	43	245.2	30.2	3125	5	I15009	I15009 Sequence 19
	44	233.6	28.7	4664	5	XLAJ2493	AJ002493 Xenopus 1
c	45	230.2	28.3	63166	55	AC023783	AC023783 Homo sapl

ALIGNMENTS

```
RESULT 1

LOCUS       A28003      3132 bp    DNA            PAT       25-SEP-1995
DEFINITION  H.sapiens HEK gene.
ACCESSION   A28003
VERSION     A28003.1  GI:1247486
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3132)
  AUTHORS   .
  TITLE     A NOVEL RECEPTOR-TYPE TYROSINE KINASE AND USE THEREOF
  JOURNAL   Patent: WO 9300425-A 7 07-JAN-1993;
FEATURES             Location/Qualifiers
     source          1..3132
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            100..3051
                     /gene="HEK"
     CDS             100..3051
                     /gene="HEK"
                     /codon_start=1
                     /protein_id="CAA01906.1"
                     /db_xref="GI:1247487"
                     /db_xref="SWISS-PROT:P29320"
                     /translation="MDCQLSILLLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDR
                     ILKLNTEIREVGPVNKGFYLAPQDVGACVALVSVRVYYKKCPFTVKNLAMFPDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
                     FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKPPSSMACTRPPSSPRNVIS
                     NINETSIVILDWSWPLDTGGRKDVTFNIICKKCGWNIKQCEPCSPNVRFLPRQFGLNVT
                     TVTVDLLAHTNYTFEIDAVNGVSELSSPPRQFAAVSITTNQAAPSPVLTIKKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEDSYTILRARGTNVTISSLRPDTIYVLQ
                     IRARTAAGYGINSRKEFETFSPDSFSISGESSQVVMIAISAAVAILLTVVIYVLCR
                     FCGYKSKHGADEKRLHFGNGHLKLPGLRTYVDPHTYEDPNQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSRPVMIVTEYMENGSLDSFLRKHDAQFTVIQLLGMLRGIASGMKYLSDM
                     GYVHRDLAARNILLINNLVCKVSDFGLSRVLEDDPEAAYTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIRAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDISTER
                     TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
                     ETQSKNGPVPV"

BASE COUNT      888 a      709 c      761 g      774 t
ORIGIN

Query Match             100.0%;  Score 813;  DB 5;  Length 3132;
Best Local Similarity   100.0%;  Pred. No. 4.3e-241;
Matches 813;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 atggattgtcagctcagctccatcctcctcctcttctcctgcagcgtgcttcggg   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 atggattgtcagctcagctccatcctcctcctcttctcctgcagcgtgcttcggg  159

Qy     61 gaactgattccgagccttccaatgaagtcaatcctggataccggagacatcaaggg  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 gaactgattccgcagccttccaatgaagtcaatctactggataccggagacatcaaggg  219

Qy    121 gagctgggctggatctcttatccatccatcacatgggtgggaagagaatcagtgg  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 gagctgggctggatctcttatccatccatcacatgggtgggaagagaatcagtgg  279

Qy    181 cattacacacaccatcaggacttaccaggtgtgcaatgtcaatgaacctcaaaacaat  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 cattacacacaccatcaggacttaccaggtgtgcaatgtcaatgaacctcaaaacaat  339

Qy    241 tggctgagacaaactggtcccagaagatttatgtgggagctcaag  300
          |||||||||||||||||||||||||||||||||||||||||||||
Db    340 tggctgagacaaactggtcccagaagatctcagaagaactcagctcagacaagaa  399

Qy    301 ttcactctacgagagctgcaatagcattccattggtttcattggaacttgcaagggagacatc  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400 ttcactctacgagagctgcaaatagcattccattggtttcattggaacttgcaagggagacatc  459

Qy    361 aacctgtactacatggagtctgatgatgatcatgggtgaaattctgagagcatcagttt  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 aacctgtactacatggagtctgatgatgatcatgggtgaaattctgagagcatcagttt  519

Qy    421 acaaagatgtgacaccattgcagcctgatgaaagtttcactcactcaaatgaatcctg  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520 acaaagatgtgacaccattgcagcctgatgaaagtttcactcactcaaatgaatcctg  579

Qy    481 atttctgaagctcaacacactgagatctgagagtagggatcttcaacacagagatttat  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    580 atttctgaagctcaacacactgagatctgagagtagggatcttcaacacagagatttat  639

Qy    541 ttggcatttcaagatgttggttgtcttgttgtcttgtcctttggttgtgtgagtatactttcaaa  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    640 ttggcatttcaagatgttggttgtcttgttgtcttgtcctttggttgtgtgagtatactttcaaa  699

Qy    601 aagtgcccaattacagtgaagaatctggcctatgtttcgccaccatggatcc  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    700 aagtgcccaattacagtgaagaatctggcctatgtttcgcccaccatggatcc  759

Qy    661 cagtccccgtggaggttagagggtctctgtcaacaattctaaggggaagatcctca  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 cagtccccgtggaggttagagggtctctgtcaacaattctaaggggaagatcctca  819

Qy    721 aggatgtactgcagtacagaagcgaatggctgttgtacccatggcaagtgttcctgcaat  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    820 aggatgtactgcagtacagaagcgaatggctgttgtacccatggcaagtgttcctgcaat  879

Qy    781 gctggcctatgaaagaaagaggtttctatgtgccaa  813
          ||||||||||||||||||||||||||||||||||||
Db    880 gctggcctatgaaagaaagaggtttctatgtgccaa  912

RESULT 2

LOCUS       I68018      3132 bp    DNA            PAT       04-FEB-1998
DEFINITION  Sequence 9 from patent US 5674691.
ACCESSION   I68018
VERSION     I68018.1  GI:2830140
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3132)
  AUTHORS   Boyd,A.W., Simpson,R.John, Wicks,I., Ward,L.David and Wilkinson,D.
  TITLE     Method of screening for ligands to a receptor-type tyrosine kinase
  JOURNAL   Patent: US 5674691-A 9 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1..3132
                     /organism="unknown"

BASE COUNT      888 a      709 c      761 g      774 t
ORIGIN

Query Match             100.0%;  Score 813;  DB 5;  Length 3132;
Best Local Similarity   100.0%;  Pred. No. 4.3e-241;
Matches 813;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 atggattgtcagctcagctccatcctcctcctcttctcctgcagcgtgcttcggg   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 atggattgtcagctcagctccatcctcctcctcttctcctgcagcgtgcttcggg  159

Qy     61 gaactgattccgagccttccaatctctactggatccggagacatcaaggg  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 gaactgattccgcagccttccaatgaagtcaatctactggataccggagacatcaaggg  219

Qy    121 gagctgggctggatctcttatccatcacatgggtgggaagagaatcagtgg  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT  3
LOCUS       HUMHEK      3149 bp    mRNA              PRI       31-DEC-1994
DEFINITION  Human receptor tyrosine kinase (HEK) mRNA, complete cds.
ACCESSION   M83941
VERSION     M83941.1  GI:183931
KEYWORDS    receptor protein-tyrosine kinase.
SOURCE      Homo sapiens lymphoid tumor cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3149)
  AUTHORS   Wicks,I.P., Wilkinson,D., Salvaris,E. and Boyd,A.W.
  TITLE     Molecular cloning of HEK, the gene encoding a receptor tyrosine
            kinase expressed by human lymphoid tumor cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1611-1615 (1992)
  MEDLINE   92179233
FEATURES             Location/Qualifiers
     source          1..3149
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="LK63"
                     /tissue_type="lymphoid tumor"
     gene            101..3052
                     /gene="HEK"
     CDS             101..3052
                     /gene="HEK"
                     /codon_start=1
                     /product="receptor protein kinase"
                     /protein_id="AAA56633.1"
                     /db_xref="GI:183932"
                     /translation="MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNWLRTNWVPRNSAQKIVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDR
                     ILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
                     FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKDPPSMACTRPPSSPRNVIS
                     NINETSVILDWSWPLDTGGRKDVTYNILCKKCGWNIIKQEPCSPNVRFLPRQFGLINT
                     TVTYDLLAHTNYTFEIDAVNGVSELSSPPRQFAAVSITNQAAPSPVLTIKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEQETSYTILRARGTNVTISSLKPDTIYVFQ
                     IRARTAAGYGTNSRKFEFETSPDSFSIGSGESSQVVMIAISAAVAILLIILVVIYLIGR
                     FCGYKSKHGADEKRLHFGNGHLKLPGLRTYVDPHTYEDPTQAVHEFAKELDANISID
                     KVVGAGEEGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSRPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDISTFR
                     TTGDWLNGVRTAHCKEIFTGVEYSSCDDIIAKISTDDMKKVGVTVVGPQKKIISSIKAL
                     ETQSKNGPVPV"
BASE COUNT      891 a    711 c    768 g    779 t
ORIGIN
```

Query Match 100.0%; Score 813; DB 9; Length 3149;
Best Local Similarity 100.0%; Pred. No. 4.3e-241;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgggattgtcaagtctctccatcatcctcctcttctcaagctctgtactcggg     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 ATGGGATTGTCAAGTCTCTCCATCATCCTCCTCTTCTCAAGCTCTGTACTCGGG    160

Qy     61 gaactgattccgcagcctccaatgaagtcaaatccagctcctcaaaaacaattcaaggg    120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 GAACTGATTCCGCAGCCTCCAATGAAGTCAAATCCAGCTCCTCAAAAACAATTCAAGGG    220

Qy    121 gagctgggctggatctcttatccatcacatgggtggaaagaaatcagctggttggatgaa    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 GAGCTGGGCTGGATCTCTTATCCATCACATGGGTGGAAGAAATCAGTGGTTGGATGAA    280

Qy    181 cattacacaccccatcaaggagacttaccaggtgtgcaatgtcatggacccacagtca    240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 CATTACACACCCCATCAAGGAGACTTACCAGGTGTGCAATGTCATGGACCACAGTCAAAAACAAT    340

Qy    241 tggctgagaacaaactgggtccccagagaactcagctcaggaagacattatgtggagctcaag    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 TGGCTGAGAACAAACTGGGTCCCCAGAGAACTCAGCTCAGGAAGACATTATGTGGAGCTCAAG    400

Qy    301 ttcactctacgagactgcaatagcattccattggtgttaggaacttgcaaggagacatcc    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 TTCACTCTACGAGACTGCAATAGCATTCCATTGGTTTTAGGAACTTGCAAGGAGACATTC    460

Qy    361 aacctgtactacatggagtctgatgatgatcatggggtgaaatttcgagagcatcagtt    420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 AACCTGTACTACATGGAGTCTGATGATGATCATGGGGTGAAATTTCGAGAGCATCAGTTT    520

Qy    421 acaaagattgacaccattgcagctgatgaaagttcactcaaatggatcttgggggaccgt    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ACAAAGATTGACACCATTGCAGCTGATGAAAGTTCACTCAAATGGATCTTGGGGGACCGT    580

Qy    481 attcttgaagctcaacctgaagattagagaaggtaggtcctgtcaacacaagaaggatttat    540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581 ATTCTTGAAGCTCAACACTGAGATTAGAGAAGGTAGGTCCTGTCAACACAAGAAGGATTTTAT    640

Qy    541 ttggcaatttcaagatgttggttgcttggttgtgctgtgagagtatacttcaaa    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    641 TTGGCCATTTCAAGATGTTGGTTGCTTGGTTGTGCTGTGAGAGTATACTTCAAA    700

Qy    601 aagtgcccatttacagtgaagaatctggcctatgtttccagacacacggtacccagtgactcc    660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

RESULT 4

```
LOCUS       RNU69278   3077 bp   mRNA   ROD   23-MAR-1998
DEFINITION  Rattus norvegicus eph-related receptor tyrosine kinase homolog
            (Rek4) mRNA, complete cds.
ACCESSION   U69278
VERSION     U69278.1  GI:1943913
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McLiernan,C.F. and Feldman,A.M.
  TITLE     IL-1 beta alters the expression of the receptor tyrosine kinase
            gene r-EphA3 in neonatal rat cardiomyocytes
  JOURNAL   Am. J. Physiol. 274 (1), H331-H341 (1998)
  MEDLINE   98120505
REFERENCE   2  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McLiernan,C.F. and Feldman,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1996) Cardiology, University of Pittsburgh, 200
            Lothrop Street, Pittsburgh, PA 15213, USA
COMMENT     On Apr 18, 1997 this sequence version replaced gi:1698721.
FEATURES             Location/Qualifiers
     source          1..3077
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /dev_stage="neonatal"
                     /cell_type="cardiomyocytes"
     gene            1..3077
                     /gene="Rek4"
     CDS             35..2989
                     /gene="Rek4"
                     /codon_start=1
                     /product="eph-related receptor tyrosine kinase homolog"
                     /protein_id="AAC06273.1"
                     /db_xref="GI:1698722"
                     /translation="MDCHLSILLFGCCVLSCSRELSPQPSNEVNLLDSKTIQGELGW
                     ISPSHGWEEISGVDEHYIPIRTYQVCNVWDHSQNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTGCKETNLYLYMESDDHGVKLELQHFKIDIIAADSSFTQWDLGDR
                     ILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
                     DSQLVEVRGSCVNNSKEEDPPRMCSTEGEWLVPIGKCTCNAGYEERGFICQACRPG
                     FYKALDGYAKCTCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACTRPPSAPRNVIS
                     NINETSVLDWSWPLDTGGRKDIFNIICKKCGNWRQCEPCSPNVRFLPRQLGLTNT
                     TVTVDLLAHTNYTFEIDAINGVSELSSPPROFAAVSITNQAPSPVMTIKKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEOETSYTILRARGTNVTISSLKPDTTYVFQ
                     IRARTAAGYGTNSRKFEFENSPDSFSISGENSHVVMIAISAAVAIVLVTVTVLVGR
                     FCGYHKSKHSSDERLHFGNGHLRLPGLRYIVDPHTYEDPTQAVHEFAKELDATNIAI
                     DKVVGAGEFGEVGSGRLRLPSKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPN
                     IIRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSD
                     MGVYHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAYTTRGGKIPVRWTSPEATA
                     YRKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPLPMDCPAALYQ
                     LMLDCWQKDRNNRKPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDIATF
                     HTTGDWLNGMRTAHCKEIFTGVEYSSCDFIAKISTDDMKKVGVTVVGPQKKISSIKA
                     LETQSKNGPVPV"
```

ORIGIN

Qy			
Db	701	AAGTGCCATTTACAGTGAAGAATCTGGCTATGTTCCAGACACGGTACCCATGGACTC	760
Qy	661	cagtccctggttggaggttaggagggtctgttcgtgtcaacaattctaaggaggatcctcca	720
Db	761	CAGTCCCTGGTGGAGGTTAGGAGGGTCTGTTGTCAACAAATTCTAAGGAGGATCCTCCA	820
Qy	721	aggatgtactgcagtacagaagoggaatggcttgtaccccattggcaagtgttcctgcaat	780
Db	821	AGGATGTACTGCAGTACAGAAGCGGAATGGCTTGTACCCCATTGGCAAGTGTTCCTGCAAT	880
Qy	781	gctggctatgaagaagaggtttatgtgccaa	813
Db	881	GCTGGGCTATGAAGAAGAGGTTTATGTGCCAA	913

RESULT 5

```
LOCUS       MUSMEK4SE   2032 bp   mRNA   ROD   15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (Mek4) secreted mRNA,
            complete cds.
ACCESSION   M68515
VERSION     M68515.1  GI:454828
KEYWORDS    receptor tyrosine kinase.
SOURCE      Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
```

ORIGIN

```
Query Match   83.5%;  Score 678.6;  DB 12;  Length 3077;
Best Local Similarity  89.7%;
Matches 729;  Conservative 0;  Mismatches 84;  Indels 0;  Gaps 0;
                               Pred. No. 2e-199;
```

Qy	1	atggattgtcagctctccatcctcctccttttcggctgctgtgtcctctcagctggctcggg	60
Db	35	ATGGATTGTCACCCTCATCCTCCTCATCCTGTTCGGCTGCTGCGTCCTCAGCTGCTCCAGG	94
Qy	61	gaactgattccgcagccttccaatgaagtcaatctcactggattcaaaaacaattcaaggg	120
Db	95	GAACTGAGTCCACAGCCTTCCAACGAAGTTAATCTACTAGATTCAAAAACGAATTCAAGGA	154
Qy	121	gagctgggtggatctctttacacagtgggtggagagaacaagtggtgtggatgaa	180
Db	155	GAGCTGGGCTGGATCTCCCTACCCATCACGGGTGGGAGGAAGATCAAGTGGTGTTGATGAG	214
Qy	181	cattacacacaccatcagagacttacaggttgtgcaatgtcatgggaccacagtcaaaacaat	240
Db	215	CATTACACACACCAAATCAGGATCTTCACCCATCCCACGGGTGGGGAAGACACAGTCAAAAATAAT	274
Qy	241	tggctggagaacacaaactggttccccagggaactcagctcagaagatttatgtgggctcaag	300
Db	275	TGGCCTGAGGACAACTGGGTGCCGACAAACTCAGCTCAGAAGACTCATGTGGGCTAAAG	334
Qy	301	ttcactctacagactgcaatagcattccattggtcttaggaacttgactgcaagagacattc	360
Db	335	TTCACACTGCGGGACTGTAACAGCATTCCATTGGTTTTGGGGACTGCAAGGAGACCTTT	394
Qy	361	aacctgtactacatggaagtcctgatcatggggtgaaattccgagaagacatagttt	420
Db	395	AACCTGTACTACATGGAGTCTGATGATCACGGTGTCAAAATTCCTAGACGCATCAGTTT	454
Qy	421	acaaagattgacacacattgcagctgatgaaagttcactccaaatggatcttgggaccgt	480
Db	455	ACAAAGATTGACACCATTGCCGCTGATGAGAAGTTTCACTCAAATGGATCTCGGGACCGC	514
Qy	481	atttctgaagctcaaacactgagagttagaagagtaggccctgtcaacaagaaggagatttat	540
Db	515	ATTTCTAAAACTCAACACTGAGATTAGAGAGTGGGGACCAGTCAATAAAAAGGGGTTTTAT	574
Qy	541	ttggtcatttcaagatgtttggttgccttggttgtctgtggagatataccttcaaa	600
Db	575	TTGGGCCTTTCAAGATGTTTGGTTGCTTTGGTTGTCTGTGGAGACTTTCAAA	634
Qy	601	aagtgcccatttacagtgaagaatctggctatgtctatgtttccagacacgtacccatggacctcc	660
Db	635	AAGTGCCCATTTACACTGAAGAATCTGGCTATGTTCCAGACACAGTGCCCATGGACTCC	694
Qy	661	cagtccctggtggaggttaggagggtctgttcgtgtcaacaaattctaaggaggatcctcca	720
Db	695	CAGTCCCTGGTGGAGGTTAGGGCCTCTTGTGTCAATAATGTCCAAGGAGGAGGACCCTCCC	754
Qy	721	aggatgtactgcagtacagaagcggaatggcttgtacccattggcaagtgttcctgcaat	780
Db	755	AGGATGTACTGCAGTACAGAAGGTGAATGGCTGGTACCCCATTGGCAAGTGCACCTGCAAT	814
Qy	781	gctggctatgaagaagaggtttatgtgccaa	813
Db	815	GCTGGGTATGAAGAACGAGGGTTCATATGCCAA	847

BASE COUNT   877 a   714 c   758 g   728 t

```
ORGANISM     CDNA to mRNA.
             Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 2032)
AUTHORS      Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
TITLE        Identification of a new eph-related receptor tyrosine kinase gene
             from mouse and chicken that is developmentally regulated and
             encodes at least two forms of the receptor
JOURNAL      New Biol. 3, 769-778 (1991)
MEDLINE      92031278
COMMENT      On Feb 16, 1994 this sequence version replaced gi:199121.
FEATURES             Location/Qualifiers
     source          1..2032
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     CDS             89..1702
                     /codon_start=1
                     /product="Mek4 secreted"
                     /protein_id="AAA39522.1"
                     /db_xref="GI:454829"
                     /translation="MDCHLSILVLLGCCVLSCSGELSPQPSNEVNLLDSKIIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQPTKIDTIAADESFTQMDLGDRI
                     LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMD
                     SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                     YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                     INETSVILDWSWPLDTGGRKDIITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                     VTVTDLLAHTNYTFEIDAVNGVSELSSPPROYAAVSTTTNQAAPSPVMTIKKDRISRN
                     SISLSWQEPEHPNGIILDYEVKYYQKEQEETSYTILRARGTNVTISSLKPDTTYVFQI
                     RARTAAGYGTNSRKEFETSPDCMYFSF"
     polyA_signal    2010..2015
     polyA_site      2032
BASE COUNT   578 a   473 c   460 g   521 t
ORIGIN
        1 atggattgtcagctccagcccttccatcctctcccctctccagctgtctgttcctgcaagcttcggg  60
       89 ATGGATTGTCACCTCTCCATCCTCGTCCTGCTCGGCTGCTGCGTCCTTCAGCTGCTCCGGA  148
      149 GAACTGAGCCCACAGCCTTCCAACGAAGTTAATCTACTAGATTCGAAAACAATTCAAGGA  208
      209 GAGCTCGGGTGGATCTCTTATCCATCACACGGGTGGGAAGAAGATCAGTGGTGTTGATGAA  268
      269 CATTACACACCAATCAGGACTTACCAAGTGTGCAAATGTCATGGATCACAGCCAAAATAAT  328
```

```
ORGANISM     CDNA to mRNA.
             Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RESULT   6
LOCUS        MUSMEK4      3197 bp    mRNA        ROD       15-FEB-1994
DEFINITION   Mouse eph-related receptor tyrosine kinase (Mek4) mRNA, complete
             cds.
ACCESSION    M68513
VERSION      M68513.1  GI:199119
KEYWORDS     receptor tyrosine kinase.
SOURCE       Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
             CDNA to mRNA.
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 3197)
AUTHORS      Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
TITLE        Identification of a new eph-related receptor tyrosine kinase gene
             from mouse and chicken that is developmentally regulated and
             encodes at least two forms of the receptor
JOURNAL      New Biol. 3, 769-778 (1991)
MEDLINE      92031278
FEATURES             Location/Qualifiers
     source          1..3197
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     gene            89..3040
                     /gene="Mek4"
     CDS             89..3040
                     /gene="Mek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA39521.1"
                     /db_xref="GI:199120"
                     /translation="MDCHLSILVLLGCCVLSCSGELSPQPSNEVNLLDSKIIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQPTKIDTIAADESFTQMDLGDRI
                     LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMD
                     SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                     YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                     INETSVILDWSWPLDTGGRKDIITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                     VTVTDLLAHTNYTFEIDAVNGVSELSSPPROYAAVSTTTNQAAPSPVMTIKKDRISRN
                     SISLSWQEPEHPNGIILDYEVKYYQKEQEETSYTILRARGTNVTISSLKPDTTYVFQI
                     RARTAAGYGTNSRKEFETSPDSFSISGENSHVVMIAISAAVAIILTIVVTIVLVGRF
```

Query Match   81.3%;
Best Local Similarity 89.2%;
Matches 725; Conservative 0; Mismatches 85; Indels 3; Gaps 1;
                                              Pred. No. 5.5e-194;
Score 661; DB 12; Length 2032;

```
Qy     1 atggattgtcagctccagcccttccatcctctcccctctccagctgtctgttcctgcaagcttcggg   60
          |||||||| ||  ||||||| ||||||||| ||| ||||| ||||||||| ||||||||| |||
Db    89 ATGGATTGTCACCTCTCCATCCTCGTCCTGCTCGGCTGCTGCGTCCTTCAGCTGCTCCGGA   148

Qy    61 gaactgagcccagccttccaatgaagtcaatctactggattcaaaaacaattcaaggg  120
          |||||||||| |||| |||||| ||||| |||||| ||||| ||||||||||||||| |
Db   149 GAACTGAGCCCACAGCCTTCCAACGAAGTTAATCTACTAGATTCGAAAACAATTCAAGGA   208

Qy   121 gagctcgggctggatctcttatccatcatcacacggtggtggaagaatcaagtggtgatgaa  180
          |||||||||||||||||| ||||||||||| ||||| |||| |||| ||| ||||| |||||
Db   209 GAGCTCGGGTGGATCTCTTATCCATCACACGGGTGGGAAGAAGATCAGTGGTGTTGATGAA   268

Qy   181 cattacacaccatcaggactttggcaagtgtgcaaatgtcatggaccacagtcaaaacaat  240
          |||||||||||| ||||||| || |||||||||||||||||||| |||||| |||||| ||
Db   269 CATTACACACCAATCAGGACTTACCAAGTGTGCAAATGTCATGGATCACAGCCAAAATAAT   328

Qy   241 tggctgagaacaaactgggtccccaggagaaactcagctcaagagaagatttatgtgaagctcaag  300
          |||||||||| |||||||||  ||||||||  |||  | |||| |||||| ||||| |||||||||
Db   329 TGGCTGAGGACAAACTGGGTACCCAGGAAACTCAGCTCAAGAGTTGGAGCTAAAG   388

Qy   301 ttcactctacagagctgcaatagcatcccattggttctaggtaacctgcaaggagacattc  360
          |||||| || ||||| ||||| | |||||||||||||||||| ||||||| |||||||||
Db   389 TTCACACTGCGGGACTGTAACAGCATTCCATTGGTTTTGGGGACTTGCCAAGAGACCTTT   448

Qy   361 aacctgtactacatggaggtctgatgatgatcagggtgaaatttcggagagcatcagttt  420
          ||||||||||||| |||||| || || ||||||||||||||||| ||||||||||||| ||
Db   449 AACCTGTACTACATGGAGTC---TGATGATCATGGCCGTCAAAATTCCGAGAGCATCAGTTC   505

Qy   421 acaaagattgacacctgcagctgtgaagtttcactcaaatgatcttcgggaccgt  480
          ||||||||||||||||| ||| |||| ||||| |||  ||||| ||| ||||||||| ||
Db   506 ACGAAGATTGACACCATTGCCGGTGATGAAAGTTTCACTCAGATGGATCTCGGGGATCGC   565

Qy   481 atcctgaagctcaacactgagttagcagaagctagtgtcctgtgacaacagaggaaggattttat  540
          ||||||||||||||||| |||| ||||||||||| |||||||||  ||||||||| |||||
Db   566 ATTCTGAAACTCAACACTGAGATTAGAGAAGTGGGACCAGTCAACAAAAAGGGGTTTTAT   625

Qy   541 ttggcattcaagatgttggttgcttgttgttgctctttgtctgtgagagtatactcaa  600
          ||||| |||||||| || |||| |||||| ||| | ||||||| ||| ||||| ||||||
Db   626 TTGGCCTTTCAAGATGTTGGTGCTTGTGTGGCTCTGGTGTCTGTGAGAGTGTACTTCAA   685

Qy   601 aagtgcccaatttacaagtgaagaatctggctatgtttccagacacagtaccacatggaccc  660
          ||||||||||| ||| ||||||||||  |||| |||||||||||||||| |||| || |||||
Db   686 AAGTGCCCGTTTACAGTGAAGAATCTGGCTATGTTTCCAGACACAGTGCCCATGGACTCC   745

Qy   661 cagtccctggtgagggttagaggttctggtgtcaacaattcaaggaaatcctcca  720
          |||| ||||||||| |||||| | |||||||| ||||| || ||||||| |||||||
Db   746 CAGTCTTTGGTGGAGGTTAGGGGCTCTTGTGTGAATAATTCCAAAGAGGAGGACCCTCC   805

Qy   721 aggatgtactgacgtacaagccgaatgcgctgttgtaccccattggcaagtgttcctgcaat  780
          |||||||||| |||| |||||| ||| |  |||||| |||||||||| |||||| |||||||
Db   806 AGGATGTACTGCGCACAGTGGAAGGAATGGGCTGGTCCCCATTGGCAAAATGCAAT   865

Qy   781 gctggctatgaagaagaggttttatgtgccaa  813
          |||||||| ||||||||| |||| ||| ||| ||
Db   866 GCTGGGTATGAAGAACGAGGAGGTTTCATATGCCAA  898
```

Left column

```
                CGYHKSKHSAEEKRLHFGNGHLHKLPGLRTIYDPHTIYEDPTQAVHEFAKELDATNISID
                KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
                IRLEGVVTKSKPEMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAMSY
                RKFTSASDVWSIGIVLWEVMSYGERPYSQMSNQDVIRAVDERYRLPPPMDCPAALYQL
                MLDCWQKDRNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDIATFH
                TTGDWINGMRFAHCKEITFGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISTIKAL
                ETQSKNGPVPV"

BASE COUNT      907 a     767 c     775 g     748 t
ORIGIN

Query Match       81.3%;     Score 661;     DB 12;     Length 3197;
Best Local Similarity     89.2%;     Pred. No. 5.8e-194;
Matches 725;     Conservative  0;     Mismatches  85;     Indels  3;     Gaps  1;

Qy     1 atggattgtcagtctccatcctcctcctctcagcgctgcctcggg         60
         |||||||||| |||||||| ||||| ||||||||||||||||||||
Db    89 ATGGATTGTCACCTTCCATCCTCCATCCTCCCTCCTCTCAGCGCTGCCTCGGA     148

Qy    61 gaactgattcccgagcctccatcctcctccctcttctcgacagcttcggg        120
         |||||| | ||| |||| |||||||||||||||||||| |||||||||||
Db   149 GAACTGAGCCCACACGCCTTCCAACGAAGTTAATCTACTTGCGTCCCCGGA       208

Qy   121 gagctgggctggatctcttatccatccatcacacatgggatgaa         180
         ||||||||| ||||| ||||||||||||||| ||||||||||||
Db   209 GAGCTGGGCTGGATCTCTTACCCATCCATCACACATGGGTCGTGTTGATGAA      268

Qy   181 cattacacacccatcaggactcattacaggtcatggacacagtcaaaacaat      240
         ||||||| || |||||||||| ||||| |||| |||||||| ||||||| ||
Db   269 CATTACACCAATCAGGACTTACCAGGTGTGCAAATGTCATTGGATCCAAAATAAT     328

Qy   241 tggctgagaacaaaactgggtcccaggaacctcagctcagcttatgtgtgagctcaag    300
         ||||||||||| |||| ||| |||||||||| |||||||| |||||||| |||||||||
Db   329 TGGCTGAGGACAAAACTGGGTACCCAGGAACCTCAGAAGACTATGTGGAGCTCAAAG     388

Qy   301 ttcaactctacgagagactagacaattccatccatggttcttaggagcaagggagacatc    360
         ||||||||| || ||||| ||||| |||||||| |||||| |||||||| ||||| |||
Db   389 TTCACACTGCGGGGACTCTGTAAACGAGACCATTCCATTGGTTTTTGGGACTTCAAGAGAC     448

Qy   361 aacctgtactacatgtcgagtctgatgatcatcgggtcgtgaaatttcgagagcatcagtt    420
         |||||||||| |||||| |||||| ||||||| |||||| ||||||||||| ||||||||||
Db   449 AACCTGTACTACATGTCCGAGTCTGATGATGATCGGGTCTGAAATTCCGAGAGCATCAGTTC     505

Qy   421 acaaagattgacacacattgacctgaaatttcactcaaatggtcttggggacgt    480
         ||| ||||| |||| ||||| ||||||||| |||||| ||| ||||||||||||
Db   506 AACCTGTACTACATGTCCGACATCCTGATGAAGTTTCACTCAAGATTCTCGGGGATCGC     565

Qy   481 atttctgaagctcaacactgagattagagaagtaggtcctgtcaacaagaaaggatttat    540
         ||||| |||| || ||||||||||| |||||||||||||||| ||||||||| ||| |||
Db   566 ATTCTGAAAACTCAACACTGAGATTAGAGAAGTGGGGACCAGTCAACAAAAAGGGTTTTAT     625

Qy   541 ttggcattcaagtgttcggtcgttctgtcctgagagtgactacttcaaa    600
         ||||||||| || ||| |||||| ||||||||||||||| |||||||||
Db   626 TTGGCCTTCAAGAGTTTCAGTGTGTTGGCTGTTGTCTGTGAGAGTGACTTCAAA      685

Qy   601 aagtgccattcacagacagtggaagaatctgctgtatgtttctaccccatgggactcc    660
         |||||||||||| |||||||||||| |||||||| |||||||| |||||| ||||||
Db   686 AAGTGCCGTTCACAGAAGAATCTGGCTATGTTTCCAGACACAGTGCCCATGGGACTCC      745

Qy   661 cagtccctggttcgaaggttcagggtccgtcaacaattctaaggagggaagatcctcca    720
         ||||||| |||| ||| ||||||||| |||||||||||||| ||| ||||| |||||||
Db   746 CAGTCTCTTGGTGGAGGTTAGGGGCTCTTGTGTCAATCAAAGAGGAGGAGACCCCC      805

Qy   721 aggagatgtactgcagtacagagaggaggcgaatggtcaagtgttcctgcaat    780
         |||||| ||||| ||||| |||||||||| ||| ||||||||||||||| |||
Db   806 AGGATGTACTGCAGCACAGAGGGATGGGAAGCTCGTGCCCATTGGCAAATGCACTTGCAAT     865

Qy   781 gctggtggctatgaagaaagaggtttcatgtgccaa    813
         |||||| |||| |||||||||||||||| ||||||
Db   866 GCTGGGTATGAAGAACGAGGTTTCATATGCCAA       898
```

Right column

```
RESULT    7

CHKCEK4
LOCUS       CHKCEK4       3241 bp    mRNA            VRT       15-FEB-1994
DEFINITION  Chicken eph-related receptor tyrosine kinase (Cek4) mRNA, complete
            cds.
ACCESSION   M68514
VERSION     M68514.1  GI:454809
KEYWORDS    receptor tyrosine kinase.
SOURCE      Gallus gallus 10 day embryo cDNA to mRNA.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 3241)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
COMMENT     On Feb 16, 1994 this sequence version replaced gi:211446.
FEATURES             Location/Qualifiers
     source          1..3241
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /dev_stage="10 day embryo"
                     /tissue_type="embryo"
     gene            32..2983
                     /gene="Cek4"
     CDS             32..2983
                     /gene="Cek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA48666.1"
                     /db_xref="GI:211447"
                     /translation="MDRRLPLLLLCAALGSAGRLSRPGNEVNLLDSKTIQGELGWI
                     SYPSHGWEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTL
                     RDCNSIPLVLGTCKETFNLYYMESDDDHLAKFREHQFTKIDTIAADESFTQMDLGDRI
                     LKINTEVREVGPYSKKGEYLAFQDVGACVALVSVRVFYKKCPFTVKNLAMFPDTVPMD
                     SQSLVEVRGSCVNHSKEEEPRKMYCSTEGEWLVPIGKCLCNAGYEERGFMCQAACRPGF
                     YKASAGNYKCAKCPPHSSTYEEASLNCRCERNYFRSEKDPPSMACTRPPSAPRNVISN
                     INETSVILDWSWPLDTGGRKDVTFNIICKKCGGSSKICEPCSDNVRFLPRQTGLTNTT
                     VTVVDLLAHTNYTFEIDAVNGVSDLSLSLRQFAAVSIITNQAAPSPIIVIRKRDRTSRN
                     SVSLSWQPEHPNGIILDYEVKYYEKQEQETSTILRAKSTNVIISGLKPDTTYVFQI
                     RARTAARVGTSSRKFEEEIETSPDSFSISSENSQVVMIATSAAVAILILVVVVLIGRF
                     CGYKKSKHGTIDKRLHFGNGHLKLPGLRTIYDPHTYEDPNQAVHEFAKELDASNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKAGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSFQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNRPKFEQIVSILDKLIRNPSSLKIITNAAARPSNLLLDQSNIDISAFR
                     TAGDWLNGFRTCQCKGIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIVSSIKTL
                     EFHTKNSPVPV"
     polyA_site      3241

BASE COUNT      916 a     735 c     794 g     796 t
ORIGIN

Query Match       67.2%;     Score 546.6;     DB 4;     Length 3241;
Best Local Similarity     80.1%;     Pred. No. 1.7e-158;
Matches 642;     Conservative  0;     Mismatches 159;     Indels  0;     Gaps  0;

Qy    13 ctctccacatcctcctcctccagcgtgcctctgttcctgcaagcttcggttggggaacttccg     72
         |||| || |||||| || |||||||| |||| ||||||||||||||||||| |||||| |||
Db    41 CGCCCCGGCCTCCGGTGCTGGCGTGCCTGGCCTGCGCCGCCCTCGGCTCCGCCGGGCGGC      100

Qy    73 cagcctccaatgaagtcaatctactggattcaaacaaaacaattcaagtggatgaacagtgg    132
         || ||| ||| ||||| |||||| ||| ||||||| ||||||||| | |||||||||| ||
Db   101 CGCCCGGGCAACGAAGTTAATCTCCTGGATTCAAAAACAATTCAAGGGGGAGCTGGGCTGG     160

Qy   133 atcctcttatccacatgggtgggaagaagatcaagtggtggatgaacattacaccacacc    192
         ||||| ||||||||| |||| |||||||| |||||||||| ||||| |||| |||| |||
Db   161 ATCTCCTACCCATCACATGGGTGGGAAGAGATTAGTGGTGTTGATGAGCATTATACTCCA      220
```

RESULT 8

```
LOCUS       I15007      3254 bp    DNA              PAT    02-APR-1996
DEFINITION  Sequence 15 from patent US 5457048.
ACCESSION   I15007
VERSION     I15007.1  GI:1249915
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3254)
AUTHORS     Pasquale,E.B. and Sajjadi,F.G.
TITLE       Eph-related tyrosine kinases, nucleotide sequences and methods of
            use
JOURNAL     Patent: US 5457048-A 15 10-OCT-1995;
FEATURES             Location/Qualifiers
source               1..3254
                     /organism="unknown"
BASE COUNT    926 a    737 c    796 g    795 t
ORIGIN
```

```
Qy  Query Match   67.0%;  Score 545;  DB 5;  Length 3254;
    Best Local Similarity   80.0%;
    Matches 641;  Conservative 0;  Mismatches 160;  Indels 0;  Gaps 0;
```

```
Qy  193 atcaggacttaccaggtgtgcaatgtcatgtgaccaacagtcaaaaacaattggctgagaaca 252
        ||||| ||||||||| |||||  ||  || ||||  ||||||||||||||||||||| |||||
Db  221 ATCAGAACTTACCAAGAGAGAGACCAATGTTATGTGGATCACAGTCAAAAACAATTGGCTGCGAACA 280

Qy  253 aactgggtcccaggaactcagatcaagtcacctacga 312
Db  281 AACTGGATTCACGAGAATTCAGCCGCAGAAGATATATGTTGGGACCTCAAGTTTACCTTGAGG 340

Qy  313 gactgcaatagcattccattggtttaggaacttcaaggagacattcaacctgtactac 372
Db  341 GACTGCAATAGTATCCCTCTAGTTCTGGGCACTTGCAAAGACTTCAAATCTGTATTAC 400

Qy  373 atgggagtctgatgatgatgggggtgaaatttcgagagagcacaagatgattgac 432
Db  401 ATGGAATCCGATGAGAGACCAATTTGGCAAAGTTCAGAGAGAGATTTATGTGGAGCTCAAGTTCA 460

Qy  433 accattgcagctgtagaaagttcactcaaatggatcttgggagccgtattctgaagctc 492
Db  461 ACCATTGCGGCTGTAGAGAGCTTCACCCAGATGGACTGGGAGCTCAAGTTTACCTTGAGG 520

Qy  493 aaacactgagattagagaagtaggtcctgtcaacacagaagggattttattggcattcaa 552
Db  521 AATACCGAAGTCCGCGAGGTGGGACCTGTTAGTAAGAAGGCTTTACTTGGCTTTCCAA 580

Qy  553 gatgttggtgctttgttgttgctgtctgtgagaagtatactttcaaaagtgcccatt 612
Db  581 GATGTAGGTCCATGGTGTTGCCCTTAGTCTCGGTGGCGACTGTACTTCAACTCCTTTC 640

Qy  613 acagtgaagaatctgctatgtttccagacacgtgtaccccatggaactcccgtg 672
Db  641 ACTGTCAAGAACCTGCCCATGTTTCCAGATACAGTTCCTATGGACTCCCAGTGTCCCTGGTG 700

Qy  673 gaggttagagggttcttgtgtcaacaattctaaggagaggagaaatcctccaaggatgtactgc 732
Db  701 GAGGTCGGGGGTTCTTGTGTCAAATCATTCAAGGAGAGGCCACCCAAGATGTACTGC 760

Qy  733 agtacagaaggcgaatgggcttgtgtacccaattggcaagtgttcctgcatgtgggctatgaa 792
Db  761 AGCACGGAAGGGAGGAATGGCTAGTGCCCATAGGGAAGTGCTTGTGTATGCTGGCTATGAA 820

Qy  793 gaaagaggtgtttatgtgccaa 813
Db  821 GAGAGAGGCTTTGCGTGCCCAA 841
```

RESULT 9

```
LOCUS       CHKCEK8     3600 bp    mRNA             VRT    01-FEB-2000
DEFINITION  Gallus gallus mRNA for Cek8, complete cds.
ACCESSION   D38174
VERSION     D38174.1  GI:1236732
KEYWORDS    receptor tyrosine kinase; Cek8.
SOURCE      Gallus gallus spinal cord cDNA to mRNA.
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (sites)
AUTHORS     Ohta,K., Nakamura,M., Hirokawa,K., Tanaka,S., Iwama,A., Suda,T.,
            Ando,M. and Tanaka,H.
TITLE       The receptor tyrosine kinase, Cek8, is transiently expressed on
            subtypes of motoneurons in the spinal cord during development
JOURNAL     Mech. Dev. 54 (1), 59-69 (1996)
MEDLINE     96404128
```

```
Db   41 CCCCGCCTGCCCTGCGCTGCTGCCCTGCGCTGTCTGGCGTGCCCTGCCCTTCGGCCTGCCCTCGGCGCTGGGGGGTCGGGGAGCGGGGTGCTGGGAGCGCCC 100

Qy   73 cagccttccaatgaactcaagtccaatctactggtgattcaaaacaaaccaattcaaagagttgagctgg 132
Db  101 CCCCCGGCCAACGAAGATTAATCTGCCTGGATTCAAAAACAAGGGGGAGGGCTGG 160

Qy  133 atctcttatcctcatcacaatgggtggggaagagatcaagttggtggttgggatgaacattacacacc 192
Db  161 ATCTCCTACCCATCACATCACAATGGGTGGGAGGAGAATTAGTGGTTGTTGGGATGAACATTACTCCA 220

Qy  193 atcaggacttaccaggtgtgcaatgtcatgtgaccaacagtcaaaaacaattggctgagaaca 252
Db  221 ATCAGAACTTACCAAGAGAGAGACCAATGTTATGTGGATCACAGTCAAAAACAATTGGCTGCGAACA 280

Qy  253 aactgggtcccaggaactcagatcaagtcacctacga 312
Db  281 AACTGGATTCACGAGAATTCAGCCGCAGAAGATATATGTTGGGACCTCAAGTTTACCTTGAGG 340

Qy  313 gactgcaatagcattccattggtttaggaacttcaaggagacattcaacctgtactac 372
Db  341 GACTGCAATAGTATCCCTCTAGTTCTGGGCACTTGCAAAGACTTCAAATCTGTATTAC 400

Qy  373 atgggagtctgatgatgatgggggtgaaatttcgagagagcacaagatgattgac 432
Db  401 ATGGAATCCGATGAGAGACCAATTTGGCAAAGTTCAGAGAGAGATTTATGTGGAGCTCAAGTTCA 460

Qy  433 accattgcagctgtagaaagttcactcaaatggatcttgggagccgtattctgaagctc 492
Db  461 ACCATGGCGGCTGTAGAGAGCTTCACCCAGATGGACTGGGAGCTCAAGTTTACCTTGAGG 520

Qy  493 aaacactgagattagagaagtaggtcctgtcaacacagaagggattttattggcattcaa 552
Db  521 AATACCGAAGTCCGCGAGGTGGGACCTGTTAGTAAGAAGGCTTTACTTGGCTTTCCAA 580

Qy  553 gatgttggtgctttgttgttgctgtctgtgagaagtatactttcaaaagtgcccatt 612
Db  581 GATGTAGGTCCATGGTGTTGCCCTTAGTCTCGGTGGCGACTGTACTTCAACTCCTTTC 640

Qy  613 acagtgaagaatctgctatgtttccagacacgtgtaccccatggaactcccgtg 672
Db  641 ACTGTCAAGAACCTGCCCATGTTTCCAGATACAGTTCCTATGGACTCCCAGTGTCCCTGGTG 700

Qy  673 gaggttagagggttcttgtgtcaacaattctaaggagaggagaaatcctccaaggatgtactgc 732
Db  701 GAGGTCGGGGGTTCTTGTGTCAAATCATTCAAGGAGAGGCCACCCAAGATGTACTGC 760

Qy  733 agtacagaaggcgaatgggcttgtgtacccaattggcaagtgttcctgcatgtgggctatgaa 792
Db  761 AGCACGGAAGGGAGGAATGGCTAGTGCCCATAGGGAAGTGCTTGTGTATGCTGGCTATGAA 820

Qy  793 gaaagaggtgtttatgtgccaa 813
Db  821 GAGAGAGGCTTTGCGTGCCCAA 841
```

```
REFERENCE    2 (bases 1 to 3600)
  AUTHORS    Ohta,K.
  JOURNAL    Unpublished (1996)
REFERENCE    3 (bases 1 to 3600)
  AUTHORS    Ohta,K.
  TITLE      Direct Submission
  JOURNAL    Submitted (09-SEP-1994) to the DDBJ/EMBL/GenBank databases.
             Kuninasa Ohta, Kumamoto University Graduate School of Medical
             Sciences, Dept. of Neuroscience and Immunology; 4-24-1 Kuhonji,
             Kumamoto, Kumamoto 862, Japan
             (E-mail:ohta9203@gpo.kumamoto-u.ac.jp, Tel:096-344-2111(ex.6754),
             Fax:096-364-3554)
FEATURES             Location/Qualifiers
     source          1..3600
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="spinal cord"
     sig_peptide     3..47
                     /codon_start=1
     CDS             3..2963
                     /codon_start=1
                     /product="Cek8"
                     /protein_id="BAA07373.1"
                     /db_xref="GI:1236733"
                     /translation="MAGVPVGALLPLLVGVGAVTGGSRVVPANEVTLLLDRSVQGELG
                     WIASPLEGGWEEVSIMDEKNPIRTYQVCNVMEPSQNNWLRTDWIPREGAQRVYIEIK
                     FTLRDCNSLPGVMGTCKETFNLYYAESNDDKERFIRESQFAKIDTIAADESFTQVDIG
                     DRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFDTI
                     TGADSSLVERGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGYEERNGECQAN
                     KIGYYKALSTDVACARCPPHSISIWEGSISTCDRGFFRAENDAASMPCTRPPSAPQN
                     LISNVNETSVNLEWSAPQNDGGRDDISYNVVCKRCRGAGESHCRCCGSGVHVSGGQNG
                     LKTTKVSITDLLAHTNYTFEVWAVNGVSKHNPSQDQAVSVTVTNQAAPSPIALIQAK
                     EITRHSVALAWLEPDRPNGVILEYEVKYYEKDQNERSYRIVRKTASRNTDIKGLNPLTS
                     YVFHVRARTAAGYGDFSGSPEFTTNTVPSPIIGDGTNPTVLLVSVAGSVLVLLLIAA
                     FVISRRSRKYSKAKQEADEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCKIE
                     KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKQRRDFLSEASIMGQFDHPNI
                     IHLEGVVTKCKPVMIITEYMENGSLDSFLRKNDGRFTVIQLVGMLRGIGSGMKYLSDM
                     SYHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSYGERPYWDMSNQDVIKAIEEGYRLPPPMDCPIALHQL
                     MLDCWQKERSDRPKFGQIVNMLDKLIRNPNSLKRGSESSRPSTALLDPSSPEFSAVV
                     SVSDWLQAIKMERYKDNFTAAGYTTLEAVVHMNQDDLARIGITAITHQNKLLSSVQAM
                     RSQMQQMHGRMVPV"
     mat_peptide     48..2960
                     /product="Cek8"
BASE COUNT      944 a    876 c    935 g    845 t
ORIGIN
Qy    22 ctcctccttcttcaagctgctctcgttcctcgacagcttcgggggaactgagttccgcagcgcttcc  81
         |||||||||||||||||||||||||||||||| |||||||||||||||||| ||| |||||
Db    27 CTCCTCCCGCTCCTCTCGGCGGGCTTCGCGGGCCCGTCACCGCGCCTTCCCGCGTGTACCCGCC  86
Qy    82 aatgaagtcaattctactggattcaaaaacaattcaaggggagctgtggctgtgtggatctcttat  141
         | |||||| ||||  |||| |||| | |||||  ||||||||||||||||||||| || |||||
Db    87 AACGAAGTGACCCTGCCTTGCAGGGGGACTCCCGCGTGCAGGGGGAGCTGGGCGCTGGATCGCCAGC  146
Qy   142 ccatcacat--gggtggggaagagaagatcaagtggtgtggatgaacattacacaccatcaagg  198
         |||||||||  ||||||||||||||||||||||||| ||||||| |||||||||||||||||
Db   147 CCGCTGGAAGGAGGAGGTGGGAAGTGAGCCATAATGGATGAGAAGAACACTCCGATCCGC  206
Qy   199 acttaccaggttgtgccaagtcaatgtccagtcaaacacaattggctgagacaaaactgg  258
         |||||||||| ||| |||||| ||||||| |||||| ||| ||| | | |||||||||
Db   207 ACCTACCAAGTTGTCAAATGTAATGGAGGCCCAGTCACAGAATAAATTGGTTACGAACTGATTGG  266
Qy   259 gtcccaggaactcaagctcagagagatttatgtggagctcaagttcactctaggagaactgc  318
         ||||||| |||| |||||||||||||||| || | |||||||||| |||| ||||| |||
Db   267 ATTCCCCCGCGAGGGGCCTCAGAGGGTATATATTGAAATCAAGTTCACGCTGAGAGACTGC  326
Qy   319 aatagaattcattgtttgaggagaacattcaaactgtactacacatggag  378
         ||| ||| ||||| || ||| ||  ||| ||  |||  | || |  ||
Db   327 AACAGCCTGCCCAGGTGTCATGGGAACCTGCCAAAGAGACTTTCAACCTCTATTACTATGAA  386
Qy   387 TCAACAGACAACAAGGACGGTTTTTATTCGAGCAGCCAGTTTGCCAGATTGACACCATT  446
         |||||||||||||||||||| |||||||||||||||||| |||||| |||||||||||
Db   447 GCTGCTGATGAGAGACGTTCACCCAGCGTGGACATTGGTGACAGATAAGGTCGAATACA  506
Qy   439 gcagctgatgaaagttcactcaaaatggatcttcttgggggaccgtattcgaagtcaaacct  498
         ||||||||||||||||||| |||||||||| |||||||||| ||||||  ||||||||||||
Db   507 GAGGTGCGGGGAGCTGTGGGCGCCTTCGACGCAGGAAGAAAAGGGTTTTACTTGGCGTTCCAGGATGTC  566
Qy   499 gagattagagaagtaggtccttgtcaacaagagaaggaggattcatttcttgggcattcaagtgtt  558
         ||||||||||||| ||| |||| ||| || ||||||  |||| || |||| |||| ||||||
Db   567 GGTGCCTGCCATTCCATTTCGTTGGTTGTTCGTTCGTCTATAAGAAGTGCCCACTGACAGTT  626
Qy   559 ggtgcttgtgttggttcctggtgttgtctgtgagagtatacttcaaaagtgcccattacagtg  618
         ||| ||||||||| || ||  |||| ||| ||||||||| ||||||||||| |||||||||
Db   627 CGAAACCTGGCACAGTTTCCAGACACCATTACTGGGGCCGGCTAAAGACATCCTCTGTGGGAA  686
Qy   619 aagaaatctggctatgttccagacacggtaccccatg--gactcccagtccctggggag  675
         ||||||||||||||||||||||||| || |||| ||  |||||| || |||| |||||
Db   687 GTTCCGTGGCCTCCTGCTGTGCAACAACTCCGAAGAGAAAGACGTGCCAAAAATGTACTGCGGG  746
Qy   736 acagaaaggcgaatggcttgtacccattggcaagtgttcctgcaatgctggctatgaagaa  795
         |||| ||||  || ||||||   |||||| ||||| || ||  |||||||| ||||||||
Db   747 GCAGATGGTGAAGTGGCTGGTACCCATTGGCAACTGTCTGTGCAATGCTGGGCTATGAAGAA  806
Qy   796 agaggttttatgtgccaa  813
         |||||||||||||||||
Db   807 CGGAATGGTGAATGCCAA  824
```

```
RESULT 10
LOCUS       MMU58332      3943 bp    mRNA            ROD       08-MAY-1997
DEFINITION  Mus musculus receptor tyrosine kinase mRNA, complete cds.
ACCESSION   U58332
VERSION     U58332.1  GI:1457960
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3943)
  AUTHORS   Lee,A.M., Navaratnam,D., Ichimiya,S., Greene,M.I. and Davis,J.G.
  TITLE     Cloning of m-ehk2 from the murine inner ear, an eph family receptor
            tyrosine kinase expressed in the developing and adult cochlea
  JOURNAL   DNA Cell Biol. 15 (10), 817-825 (1996)
  MEDLINE   97047913
REFERENCE   2 (bases 1 to 3943)
  AUTHORS   Lee,A.M., Ichimiya,S., Greene,M.I. and Davis,J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-1996) Pathology & Laboratory Medicine, University
            of Pennsylvania, 36th and Hamilton Walk, Philadelphia, PA
            19104-6082, USA
FEATURES             Location/Qualifiers
     source          1..3943
                     /organism="Mus musculus"
                     /strain="Balb/c"
                     /db_xref="taxon:10090"
     CDS             285..3392
                     /note="eph-related tyrosine kinase; similar to Rattus
                     norvegicus receptor tyrosine kinase Ehk-2, SwissProt
                     Accession Number S51605"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAB53836.1"
                     /db_xref="GI:1457961"
                     /translation="MGGCEVREFLLQFGFFLPLLTAWTGDCSHVSNQVVLLDITTVMG
                     ELGWKIYPLNGWNDAITEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIVE
```

```
Qy      796  agaggtttatgtgcca     812
             |||||  ||||
Db     1095  ATCGAGGGTTCTTGCCA   1111
```

RESULT 11

```
Db     1095  ATCGAGGGTTCTTGCCA   1111
```

LOCUS   RNEHK1   3531 bp   mRNA   ROD   17-FEB-1997
DEFINITION R.norvegicus ehk-1 mRNA.
ACCESSION   X78689
VERSION   X78689.1 GI:531543
KEYWORDS   Ehk-1 gene; receptor tyrosine kinase.
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 3531)
  AUTHORS   Taylor,V., Miescher,G.C., Pfarr,S., Honegger,P., Breitschopf,H.,
            Lassmann,H. and Steck,A.J.
  TITLE   Expression and developmental regulation of Ehk-1, a neuronal
            Elk-like receptor tyrosine kinase in brain
  JOURNAL   Neuroscience 63 (1), 163-178 (1994)
  MEDLINE   95206467
REFERENCE   2 (bases 1 to 3531)
  AUTHORS   Miescher,G.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-APR-1994) G.C. Miescher, Neurological Clinic &
            Research Dept., University Hospitals Basel, 4031 Basel, SWITZERLAND
REFERENCE   3 (bases 1 to 3531)
  AUTHORS   Maisonpierre,P.C., Barrezueta,N.X. and Yancopoulos,G.D.
  TITLE   Ehk-1 and Ehk-2: two novel members of the Eph receptor-like
            tyrosine kinase family with distinctive structures and neuronal
            expression
  JOURNAL   Oncogene 8 (12), 3277-3288 (1993)
  MEDLINE   94067777
FEATURES   Location/Qualifiers
     source   1..3531
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /dev_stage="6 weeks post-natal"
                     /tissue_type="brain"
                     /clone_lib="lambda 936501, from Stratagene USA"
     misc_feature   1..53
                     /note="putative alt. splice region (by comparison with
                     publ. seq. and PCR cDNA fragments)"
     conflict   114
                     /citation=[3]
                     /replace="c"
     conflict   334^335
                     /citation=[3]
                     /replace=""
     gene   410..3106
                     /gene="ehk-1"
     CDS   410..3106
                     /gene="ehk-1"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="CAA55357.1"
                     /db_xref="GI:531544"
                     /db_xref="SWISS-PROT:P54757"
                     /translation="MRGSGPRGADADGPRAQASWCHARRGDTPRVPASLACCYSAPLK
                     GPLWTCLLCAALRTLLASPSNEVNLLDSRPVLGDLGWIAFPKNGWEEIGEVDENYAP
                     IHTQVCKVMEQNQNRWLLISWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETNM
                     YFEESDDENGRNIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDVGPLSKKGFY
                     LAFQDVGACIALVSVRVYYKKCPSVVRHLAVFPDITGADSSQLLEVSGSCVNHSVTD
                     DPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQVCRPGFFKASPHSQTCSKCPHSY
                     THEEASTSCVCEKDYFRRESDPPTMACTTPSPVTNVKRGIAIARKNSISLSWPDRPNG
                     IILEYEIKYFEKDQETSYTIIKSKETTITAEGLKPASVYVFQIRARTAAGYGVFSRRF
                     EFETIVFPAASNDQSOIPIIAVSTVGVILLAVMIGFLLSGSSCEGCGGRASSLCAVA
                     HPSLIWRCGYSKAKQDPEEEKMHFHHGHIKLPGVRTYIDPHTYEDPTQAVHEFAKEIE
                     ASCITIERVIGAGEFGEVCSGRLKLPGKRELPVAIKTLKVGYTEKQRRDFLSEASIMG
                     QFDHPNIIHLEGVVIKSKPVMIVTEYMENGSLDTFLKKNDGOFTVIQLVGMLRGIAAG
```

```
Qy     79  tccaatgaagtcaatactactggattcaaaaacaattcaaggggggagctgggctggatctct  138
           ||||| |||| ||||||| ||||||||||||||||||||||||| ||| |  ||||| |||
Db    375  TCCAACCAAGTTGTGTTGCTTGATACAACCACACAGTGATGGGAGAACTAGGGAAAACA       434

Qy    139  tatccatcacatgggtggaagagatcagtggtgtggatgaacattacacaccatcagg      198
           |||||| | |||| || |||||  |||  ||||||||| || |||||| |||| ||||
Db    435  TATCCATTAAATGGGTGGGGGATGCCAATACTGAAATGGATGAACACAATAGGCCCATACAT  494

Qy    199  acttaccaggtgtgcaagtgtcatggaccacagtcaaaaacaaattggtgctgagaacactgg  258
           |||| || ||||| | |||| || |||||||||||||||||| ||| |||||| |||||||
Db    495  ACATACCAGGTATGCAAGTGTCATGGACCACAAAACCAGAACAACTGGCTTCGTACTAACTGG  554

Qy    259  gtcccaggaactcagctcagaagattatactgtgagctcaagttcactctacgagactgc     318
           |||||  | |||||||||||||||||| ||||| ||||||||||||||||||| ||||||
Db    555  ATCTCTCGTGATGCTCCACAGAAAAATTATGTGGAAATGAAGTTCACACTGAGGGATTGT      614

Qy    319  aatagcattccattggttttaggaacttcaaggagagacattcaaacctgtactacatggag  378
           ||| || |||||||  ||| || ||||||||||||||||||| |||| ||| |||| ||||
Db    615  AACAGCATCCCATGGGTCTTGGGACTTGTAAAGAAAACCTTAACCTGTATTATATAGAA      674

Qy    379  tctgatgatcatgggggtgaaattcggagagcatcagtttcacaaagattgacaccatt      438
           |||||||| || ||||| ||||| ||||| || | |||||||||| |||||||||||
Db    675  TCCGGATGAATCCCATGGGAACAAAATTCAAGCCACCAATATAATAAAGATCGACACAAATT   734

Qy    439  gcagctgatgaaagtttcactccaaatggatctcttgggggacctgtattctgaagctcaacact  498
           ||||| || ||||| ||||||||||  ||| || |||||||||| || ||| ||||||||
Db    735  GCTGCTGATGAGAGTTTTACTCAGAGATGGGATTTTGGGACCGCCATCTTAAACTCAACACT  794

Qy    499  gagattagaagaagtaggtcctgtcaacaagaagagagattttattggccttcaaagatgtt     558
           ||| | || ||||  ||| ||| || |||||| || || |||||| |||| |||| |||||
Db    795  GAAAATCCGTCGAGTGGGCCCTATAGAAGAGAAAAGGATTTTATTTGGCTTTCAAGATATT     854

Qy    559  ggtgtcttgttgttgcttcctgtgaggagtatacttcaaaaagtgcccatttacagtg     618
           || ||  |||| |||||| ||||| ||||| |||  |||||||||| || ||||||||
Db    855  GGAGACATGCAATGCGCTCGGGTCTCCGAGTTTTCACAAAAAATGCCCCTTCACCGTG      914

Qy    619  aagaatctggctatgttccagacacggtaccca---tggactcccagtccctggtggag     675
           |||| ||||| ||| |||| |||| || ||||     ||| ||||| || |  ||||| 
Db    915  CGGAGCTTGGCTATGTTTCTGATACCATCCCAAGGGGGTTGATTCTCCTCTTTGGTTGAA     974

Qy    676  gttagagaggtcttgtcaacaattctaaggaggagatgactgaagt     735
           || ||| ||||| || ||  | | || ||||| ||| ||| ||| 
Db    975  GTGCGCGGGCTCATGTGTAAAGAGTGCTGAGAGCGAGAATACTCACTACTCTGGA     1034

Qy    736  acagaaaggcgaatgggctgttcctgccgatgctggctcatgaagaa     795
           || ||| ||| ||||||  || |||| | |||| || |||  ||| 
Db   1035  GCTGATGGGGATTGGCCTTGTTCCTCTCTTGGAAGGTGTATCTGCAGTACAGGGTGATGAAGAA  1094
```

```
BASE COUNT    1082 a    929 c    957 g    975 t
ORIGIN
```

```
Qy      MKFTLRDCNSIPWVLGTKETFNLYIYIESDESHGTRFKPSQYIKIDTIAADESFTQMD
        LGDRILKLNTEIREVGPIERKGFYLAFQDIGACIALVSVRVFYKKCPFTVRSLAMFPD
        TIPRVDSSSLVEVRGSCVKSAEERDTPKLYCGADGDWLVPLGRCICSTGYEEIGSCH
        ACRPGFYKAFAGNTKCSKCPPHSSTVYEEATSVCHCEKGYFRAEKDPRMSMACTRPSAP
        RNVAFNINEFALILEWSPSDTGGRRDLTVSVICKKCGLDTQCEDCGGLRFIPRHT
        GLINNSVVLDFVSHVNYTFEIEAMNGVSELSISPKPFFAITVTHDHAPSLIGMMRK
        DWASQNSLALSWQPAFSNGAILDYETKYYKEHEQLTYSSRSKAPSVIVTGLKPAT
        TYIFIRVRATIGYSGYSQKFEFETGDETSDMAAEDQGIQLIVIATAAVGGTLLVILTL
        FFLITGRCQWYIKAKMKSEEKRRIHLQNGHLRFPGIKTYIDPDTYEDPSLAVHEFAKE
        IDPSRIRIERVIGAGEFGEVCSGRLKTPGKREIPVAIKTLKGGHDRQRDFLREASI
        MGQFDHPNIIRLEGVVTKRSFPAIGVEAFCPSFLRAGFLNGIQAHPVTAGGSLPPRI
        PAGRPVMIVEYMENGSLDSFLRKHDGHFTVIQLVGMLRGIASGMKYLSDMGYVHRDL
        AARNILVNSNLVCKVSDFGLSRVLEDDPEAAYTTTGGKIPIRWTAPEAIAYRKFSSAS
        DAWSYGIVMWEVMSYGERPYWEMSNODVILSIEEGYRLPAPMGCPPSLQQLMLHCWQK
        ERNHRPKFTDIVSFLDKLIRNPSALHTLVEDILVWPESPGDVPEYPLFVTVGDWLDSI
        KMGQYKSNFWAAGFTTFDLISRMSIDDIRRIGVILIGHQRRIVSSIQTLRLHMMHIQE
        KGFHV"
```

```
                  MKYLSDMGYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWT
                  APEAIAFRKFTSASDVWSYGIVMWEVVSYGERPYWEMTNQQVIKAVEEGYRLPSPMDC
                  PAALYQLMLDCWQKDRNSRPKFDDIVNMLDKLIRNPSSLKTLVNASSRVSTLLAEHGS
                  LGSGAYRSVGEWLEAIKMGRTEIPFMENGYYSSMDAVAQVTLE"
conflict          424
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
conflict          435^436
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
conflict          458
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
conflict          462..482
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
misc_feature      934
                  /gene="ekh-1"
                  /note="putative alt. splice region (by comparison with
                  publ. seq. and PCR cDNA fragments)"
conflict          1531
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="c"
conflict          1816
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
conflict          1821
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="c"
conflict          1885..1950
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
misc_feature      2094
                  /gene="ekh-1"
                  /note="putative alt. splice region (by comparison with
                  publ. seq. and PCR cDNA fragments)"
conflict          2211
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="g"
conflict          3219^3220
                  /gene="ekh-1"
                  /citation=[3]
                  /replace="c"
conflict          3308
                  /gene="ekh-1"
                  /citation=[3]
                  /replace=""
BASE COUNT        1001 a       772 c       892 g       866 t
ORIGIN
```

```
Query Match        43.2%;   Score 351.4;  DB 12;   Length 3531;
Best Local Similarity   67.3%;    Pred. No. 5.7e-98;
Matches  512;  Conservative   0;  Mismatches  246;  Indels   3;   Gaps   1;

Qy       56 tcggggaactgattccgcagcctccaatgaagtcaatcctactggattcaaaaacaattc 115
            || |||||  |||| ||| |||| || | | || ||| | || |||| ||||| | |||
Db      579 TCCGGACCCTTTTGGCCAGCCCCAGTAACGAACTGAATTTATTGGATTCCGCGCACTGTCC 638

Qy      116 aaggggagctgggctggatctcttattccatcacacatgggtcggaagagatcagtggtg 175
            |||||  ||| ||| ||||| ||||| ||| || |  || ||| ||| ||| |||| |||
Db      639 TGGGAGACCTTGGATGGATTGCTTTTCCAAAGAATGGGTGGGAAGAGAATATCAAAGAGAA 698

Qy      176 atgaacattacacaccatcaggaacttaccaggtgttgcaaatgtcatggggaccacgtcaa 235
            |  |||| ||| || ||| | |||  ||| | |||| | ||  |||| |||  || |||| |
Db      699 ATGAAAAACTATGCCCCATCCACACCTATCCAACAAGTGCGCAAAGTTATGGGAACAGAATCAGA 758
```

```
RESULT 12
AR062744

LOCUS       AR062744     3906 bp    DNA         PAT       29-SEP-1999
DEFINITION  Sequence 102 from patent US 5843749.
ACCESSION   AR062744
VERSION     AR062744.1  GI:5990435
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3906)
  AUTHORS   Maisonpierre,P.C.; Maslakowski,P. and Yancopoulos,G.D.
  TITLE     Ehk and Ror tyrosine kinases
  JOURNAL   Patent: US 5843749-A 102 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1..3906
                     /organism="unknown"
BASE COUNT        1097 a       876 c       990 g       943 t
ORIGIN
```

```
Qy      236 acaattggctggagaacaaactgggtcccagggaacctcagcctcaagaagacgattatgtgtgggagc 295
            | |||||||||| || || |||||||| ||||| ||| |||| ||||| |||| ||| ||| |||||||||
Db      759 ATAATTGGCTGTGTTGACCACAGTTGACATCTCTAAGCAAGGTGCTTCCCAGAATTTTTATTGAAC 818

Qy      296 tcaagttcacctctacgagactgcaatagacattccattggtcttttaggaacttgcaaggaga 355
            ||||| ||| | || || ||| |||  || || || || ||||| || |  ||||| || || |
Db      819 TCAAGTTTACTCTGAGGGATTGCAACACCCTTCCTGAGGACTGGGGAGACTTGCAAGGAGA 878

Qy      356 cattcaaacctgtactacatgggagtctgatgatgatgaatggggtggaagcatc 415
            ||||||| || || || || |||||| ||||| |||||| ||||| ||| | ||
Db      879 CCTTTAACCATGTATATTATTTGAGTCGGATGAGAGGAATATCAAAGGAAGAACC 938

Qy      416 agttacaaagaattgacaccatccagtgatgataaagtttcactcaaatggatcttggg 475
            | |||| ||  || || ||  || || ||||| | |  || ||||||| |||  |||||
Db      939 AGTACATCAAAGATCGATAACCATTGCTGCTGATGGAGAGACCTTCACCGAGAGAGTTGACCTTGGAG 998

Qy      476 accggtattctgaagctccaacactgagattagaggaagtaggtcctgtcaacagaagggat 535
            || ||||| ||| ||||| || ||||| ||| || || || ||| || || |||| ||||||
Db      999 ACCGGGTCATGAAGCTGAATACGGAGGGTCAGACAGTGTTAGGAGAGACCTTCACCGAGAGAGTTTGACCTTGGAT 1058

Qy      536 tttatttggcaattccaaggatgttggttgctcttgttcgtgaagtatact 595
            || ||||||||||| || || |||| | ||||| || ||||| |  || ||
Db     1059 TTTATCTTGCTTTCCAAGATGTCGCTGGCTTGGCTTGTCCTTGTTTCGTGTCCGGTGTCTCACT 1118

Qy      596 tcaaaaaagtgcccattacagtcgaagaatctggctatgttttccagacacggttacccatgg 655
            |||| |||||| ||||| | || | || |||| ||| ||||||| ||| |  || | ||  |||
Db     1119 ATAAAAAAATGTCTTCTGTAGTTAGACAATTTGGCTGTGTTTTCCGTGGTGTTTCCCTGACAGCTGTGGAG 1178

Qy      656 actccagtccc---tggtgtgaggttagagggtcttgtgtcaacaatcttaaggaggaag 712
            ||||||||||    || ||| || | || ||  |||| | |||| |||| || ||| |||
Db     1179 CAGATTCTTCCCCCAGTTGCTACAGACGTGTCAGGCTCTGTGGTTCCCAACCATTGTGTGACAGGACG 1238

Qy      713 atcctccaaggatgtactgcagtacagaggaaggcgaatggccttgtacccaattggccaagtgtt 772
            || ||||| || ||| | |||||| ||||  ||| | |||  ||||| | ||| ||||| || |
Db     1239 ATCCTCCAAAAATGCATTGCAGTGTCGAGAGCTGCTGGAAGCTGCTGGTTCCCAACCATTTCTGTGACAGATCGACG 1298

Qy      773 cctgcaagtcgctggctatgagaagaaggtttatgtgccaa 813
            ||| |||| || ||  |||| |||  || | |||||| |||
Db     1299 TGTGCAAGCGCCGGATATGAGAAGAGAAAATGGTACCTGTCAA 1339
```

```
Query Match        43.2%;   Score 351.4;  DB 5;   Length 3906;
Best Local Similarity   67.3%;    Pred. No. 5.8e-98;
Matches  512;  Conservative   0;  Mismatches  246;  Indels   3;   Gaps   1;

Qy       56 tcggggaactgattccgcagcctccaatgaagtcaatcctactggattcaaaaacaattc 115
            || |||||  |||| ||| |||| || | | || ||| | || |||| ||||| | |||
Db      630 TCCGGACCCTTTTGGCCAGCCCCAGTAACGAACTGAATTTATTGGATTCCGCGCACTGTCC 689

Qy      116 aaggggagctgggctggatctcttattccatcacacatgggtcggaagagatcagtggtg 175
            |||||  ||| ||| |||||| ||||| ||| || |  || ||| ||| ||| |||| |||
Db      116 ATGAAAAACTATGCCCCATCCACACCTATCCAACAAGTGCGCAAAGTTATGGGAACAGAATCAGA
```

RESULT 13

LOCUS   AR025488   4165 bp   DNA   PAT   05-DEC-1998
DEFINITION Sequence 1 from patent US 5798448.
ACCESSION   AR025488
VERSION   AR025488.1   GI:3978116
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 4165)
  AUTHORS   Caras,I.W. and Winslow,J.W.
  TITLE   AL-1 neurotrophic factor antibodies
  JOURNAL   Patent: US 5798448-A 1 25-AUG-1998;
FEATURES   Location/Qualifiers
     source   1..4165
                     /organism="unknown"
BASE COUNT   1157 a   907 c   1035 g   1066 t
ORIGIN

Query Match   43.2%; Score 351.4; DB 5; Length 4165;
Best Local Similarity   67.3%; Pred. No. 5.8e-98;
Matches 512; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy   56 tcgggaactgattccgcagcctccaatgaagtcaatctactggattcaaaaacaattc 115
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   690 TGGGAGACCTTGGATTGCATTGCTTTTCCAAAGAATTGGGTGGGGAGAGATTGGTGAAGTTG 749

Qy   176 atgaacattacacacccatcaggacttacaggtgtgcaatgtcatggaccacagtcaa 235
Db   750 ATGAAAACTATCGCCCACACACACCCTTATCACAAGTGTCCAAAGTTATGGAACAGAATCAGA 809

Qy   236 acaaatgggctcaagaaaaaactgggtcccccagagaacaattcaagagaagattttatg 295
Db   810 ATAATTGGCTGTTGACCAGTTGGATCTCTACCAGAGAGATTTTTATTGGAC 869

Qy   296 tcaaagttcacctcatacgagactgcaatagacattccattggttttaggaacttgcaa 355
Db   870 TCAAGTTTTACTCTGAGGGATTCCAAACGCCTTCCTGAGGGACTTCCAAGGAGA 929

Qy   356 cattcaacctgtactacatggtatgtgatgaatctcagaagagacatc 415
Db   930 CCTTTAACATGTATAATTTTTGAGTCGGATGAGAATGGGAGAAATATCAAACAGAACC 989

Qy   416 agtttacaaaaagattgacaccattggcagctgatgaaatttcactcaaatgatcttgggg 475
Db   990 AGTACATCAAGATCGATACCAATCCATTGCTGCGTGATGATGAGAGCTTGACCTTGGAG 1049

Qy   476 accgtattctgaagctcaaacactgagattagaagaagttagtgagagaaatggatcttgggg 535
Db   1050 ACCGGGTCATGAACGTGAAATACGGGAGGTCAGAGATGTAGGAACCTCTGAGCAAAAGGGAT 1109

Qy   536 tttattggcattccaagatgttggtgtctgtgagagtatact 595
Db   1110 TTTATCTTGCTTTCCAAGATGTCGGGCCTTGCATCGGCTCTTGTCCGTGTCTACT 1169

Qy   596 tcaaaaagtgcccattacacagtgaaaaatctggctatgttccagacaacgtgtaccatgg 655
Db   1170 ATAAAAAAATGTCCTTCTGTAGTTAGACATTTGGCGTGTTTCCCTGACACGATCACTGGAG 1229

Qy   656 actcccagtccc---tggtggaggttagagggtctcttgtgtcaacaattctaaggaggaag 712
Db   1230 CAGATTCTTCCCAGTGTCCTAGAGGTGTCAGGCGTCTTCCTCCGTGTCCTAACCATTCTGTGTGACGACG 1289

Qy   713 atcctccaaggatgtactgcagtacagaaggcgaaatggcttgttacccccattggcaagtgtt 772
Db   1290 ATCCTCCCAAAATGCATTGCAGTGCTCGAAGGGGGAGTGTGGCTCCCATCCATCGTGACAGACG 1349

Qy   773 cctgcaatgctgaagaaaagaaaggagggtttatgtgccaa 813
Db   1350 TGTGCAAGGCCGGATATGAAGAGAAAAAATGGTACCTGTCAA 1390

RESULT 14

LOCUS   MMSEK   4242 bp   mRNA   ROD   30-JUN-1993
DEFINITION M.musculus mRNA for tyrosine kinase.
ACCESSION   X65138 S51422
VERSION   X65138.1   GI:54083
KEYWORDS   Sek gene; Tyrosine kinase.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 4242)
  AUTHORS   Charnay,P.
  TITLE   Direct Submission
  JOURNAL   Submitted (26-MAR-1992) P. Charnay, Ecole Normale Superieure, Lab
            de Genetique Moleculaire, URA 1302, 46 Rue d'Ulm, 75230 Paris Cedex
            05, FRANCE
REFERENCE   2 (bases 1 to 4242)
  AUTHORS   Gilardi-Hebenstreit,P., Nieto,M.A., Frain,M., Mattei,M.G.,
            Chester,A., Wilkinson,D.G. and Charnay,P.
  TITLE   An Eph-related receptor protein tyrosine kinase gene segmentally

Db   695 TCCGGACCCTTTGGCCAGCCCCAGTAACGAACGTGAATTTTATTGGATTCGCGCGACTGTCC 754

Qy   116 aaggggagctgggcgtggatcttctccatcacatcagtcgtggaagaagaatcacgtggttggg 175
Db   755 TGGGAGACCTTGGATTGCATTGCTTTTCCAAAGAATTGGGTGGGGAGAGATTGGTGAAGTTG 814

Qy   176 atgaacattacacacccatcaggacttacaggtgtgcaatgtcatggaccacagtcaa 235
Db   815 ATGAAAACTATCGCCCACACACACCCTTATCACAAGTGTCCAAAGTTATGGAACAGAATCAGA 874

Qy   236 acaaatgggctgaacaaaactgggtcccccagagaacaattcaagagaagattttatg 295
Db   875 TCAAGTTTTACTCTGAGGGATTCCAAACGCCTTCCTGAGGGACTTCCAAGGAGA 934

Qy   296 tcaaagttcacctcatacgagactgcaatagacattccattggttttaggaacttgcaa 355
Db   935 TCAAGTTTTACTCTGAGGGATTCCAAACGCCTTCCTGAGGGACTTCCAAGGAGA 994

Qy   356 cattcaacctgtactacatggtatgtgatgaatctcagaagagacatc 415
Db   995 CCTTTAACATGTATAATTTTTGAGTCGGATGAGAATGGGAGAAATATCAAACAGAACC 1054

Qy   416 agtttacaaaaagattgacaccattggcagctgatgaaatttcactcaaatgatcttgggg 475
Db   1055 AGTACATCAAGATCGATACCAATCCATTGCTGCGTGATGATGAGAGCTTGACCTTGGAG 1114

Qy   476 accgtattctgaagctcaaacactgagattagaagaagttagtgagagaaatggatcttgggg 535
Db   1115 ACCGGGTCATGAACGTGAAATACGGGAGGTCAGAGATGTAGGAACCTCTGAGCAAAAGGGAT 1174

Qy   536 tttattggcattccaagatgttggtgtctgtgagagtatact 595
Db   1175 TTTATCTTGCTTTCCAAGATGTCGGGCCTTGCATCGGCTCTTGTCCGTGTCTACT 1234

Qy   596 tcaaaaagtgcccattacacagtgaaaaatctggctatgttccagacaacgtgtaccatgg 655
Db   1235 ATAAAAAAATGTCCTTCTGTAGTTAGACATTTGGCGTGTTTCCCTGACACGATCACTGGAG 1294

Qy   656 actcccagtccc---tggtggaggttagagggtctcttgtgtcaacaattctaaggaggaag 712
Db   1295 CAGATTCTTCCCAGTGTCCTAGAGGTGTCAGGCGTCTTCCTCCGTGTCCTAACCATTCTGTGTGACGACG 1354

Qy   713 atcctccaaggatgtactgcagtacagaaggcgaaatggcttgttacccccattggcaagtgtt 772
Db   1355 ATCCTCCCAAAATGCATTGCAGTGCTCGAAGGGGGAGTGTGGCTCCCATCCATCGTGACAGACG 1414

Qy   773 cctgcaatgctgaagaaaagaaaggagggtttatgtgccaa 813
Db   1415 TGTGCAAGGCCGGATATGAAGAGAAAAAATGGTACCTGTCAA 1455

```
JOURNAL     Oncogene 7 (12), 2499-2506 (1992)
MEDLINE     93096484
REMARK      Erratum:[published erratum appears in Oncogene 1993
            Apr;8(4):1103]
FEATURES             Location/Qualifiers
     source          1..4242
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /dev_stage="8.5 days"
                     /map="band C"
     gene            58..3018
                     /gene="Sek"
     CDS             58..3018
                     /gene="Sek"
                     /note="Eph receptor related"
                     /codon_start=1
                     /product="tyrosine kinase"
                     /protein_id="CAA46268.1"
                     /db_xref="GI:54084"
                     /db_xref="MGD:MGI:98277"
                     /db_xref="SWISS-PROT:Q03137"
                     /translation="MAGIFYFILFSFLFGICDAVTSGRVYIPANEVLLLDSRSVQGELG
                     WIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRTDWITREGAQRVYIEIG
                     FTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRSEQGFKTDTIAADESPTQVDIG
                     DRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTI
                     TGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEENGGECQAC
                     KIGYYKALSTDASCAKCPPHSYSVWEGATSCTCDRGFFRADNDAAMPCTRPPSAPLN
                     LISNVNETSVNLEWSSPQNTGGRQDISYNVVCKKCGAGDPSKCRPCGSGVHYTPQQNG
                     LKTTSVTDLLAHTNYTFEIWAVNGVSKYNPSDQSVSVTVTNQAAPSSIALVQAK
                     EVTRYSVALAWLEPDRNGVLILEYEVKYYEKDQNERSYRIVRTAARNTDIKGLNPLTS
                     VVFHVRARTAAGYGDFSEPLEVTTNTVPSRIIGDGANSTVLLVSVSGSVVLVVILLAA
                     FVISRRSRKYSRAKQEADEEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCIKIE
                     KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKQRRDFLSEASIMGQFDHHNI
                     IHLEGVVTKCKPVMIITEYMENGSLDAFLRKNDGRFTVIQLVGMLRGIASGMKYLSDM
                     SVHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSTGERPYWDMSNQDVIRAIEEGYRLPPPMDCPIALHQL
                     MLDCWQKERSDRPKFGQIVNMLDKLIRNPNSLKRTSSESSRPNTALLDPSSPEFSAVV
                     SVGDWLQAIKMDRYKDNFTAAGYTTLEAVHRMSQDDLARIGITAITHQNKILSSVQAM
                     RTQMQQMHGRMVPV"
     sig_peptide     58..102
                     /gene="Sek"
     mat_peptide     103..3015
                     /gene="Sek"
                     /product="tyrosine kinase"
     misc_feature    2392..2394
                     /gene="Sek"
                     /note="tyrosine autophosphorylated"
BASE COUNT     1155 a   1034 c   1052 g   1001 t
ORIGIN
```

```
Query Match    43.2%;  Score 351.2;  DB 12;  Length 4242;
Best Local Similarity   66.5%;  Pred. No. 6.7e-98;
Matches 535;  Conservative   0;  Mismatches 263;  Indels   6;  Gaps   2;

Qy     16  tccatcctcctcctccagctgctctgttcctcgacagcttcggggaactgatccgcag   75
           |||  ||||||| |||||||||||||||||||||||||||||||||||||| ||||
Db     76  TTCATCCTCTTTCGTTTCTTTGGAATTCGCAGCCTGTCACCGGTTCACGGTGTATAC  135

Qy     76  ccctccaatgaagtcaatcctactggattcaaaaacaattcaagggaggctgggatc  135
           ||||||| | |||||||||||||||||||||||||| |||||||||||||||||||
Db    136  CCCGCGAATGAGTTACTTATTGGATTCCAGATCTGTTCAGGGAGGAGACCTTGGGTGGATA  195

Qy    136  tcttatccat--cacgtgggtgggagagatcagtggtgtggatgaacattacacaccc  192
           ||| |||| ||   |||||| |||||||||||||||||||||||||||||||||
Db    196  GCAAGCCCTCTCGGAAGGAGAGGTGGGAGGAAGTGAGCATTATGCATGGAGAAAATACACCG  255

Qy    193  atcaggaattaccaggtgtgtcaatgtcatgggaccacagtcaaaacaattggctgagaaca  252
           ||| ||| |||||||||||| |||| || ||||||||||| ||||||| |||||| ||
Db    256  ATCCGAACTACCACAGGTGTGCAAACGTGATGGAAGCCAGCGACAACAACTGTGGCTGCGAACT  315

Qy    253  aactgggtcccagggaactcagccaagaagatttatgtgggagctcaagttcactctagga  312
           ||||| |||| ||| ||||||| ||| |||| |||  | ||||| |||| |||||||||| |
Db    316  GACTGGATCACCCCGCGAAGGGGCGGCAGGAGGGGTGTACATTGAAATTAATTAAGTTCACTCTAAGG  375

Qy    313  gactgcaatagcatccattgttcttaggaaacttgcaagggaggacattcaacctgtactac  372
           |||||||| | ||| |||  ||| | || ||| ||||||| || | ||  |||| ||||||
Db    376  GACTGCAACAGTCTTCCGGCGCTCATTGGGGACTTGCAAGGAGACGTTTAACCTCTACTAC  435

Qy    373  atggagtctgatgatcatgggtgaaattcgagaggcatcaagttacaaagattgac  432
           |||||||| |||||||| ||| | | |||||||||| |||| || |||| |||||
Db    436  TATGAGTCGGACAACAGCAAGGAGACGCCTTCATCAGACGAGAGAGAGAGCGAGCTGAC  495

Qy    433  accaattgcagctgatgaaagtttccactcaaaatggaatggatcttgggggaccgtattctgaagtc  492
           ||||| |||| ||||||||||| |||| | || ||| ||| || ||  |||||||| | ||||||
Db    496  ACCAATCGCAGCCTGATGAGAAGTTTCACGCAGGTGGACAATTGGTGACCGGAATCATGAAACTC  555

Qy    493  aaacactgagcaattagaagatagtgttcctgtcaacaagaagggattttttgtgcatttcaa  552
           ||||||||| | || |||||| ||||| |||||||| |||||||||||||  | || |||| ||
Db    556  AACACTGAGATCCGGGACGTTAGGCCCACTGAGCAAGGAAAGGGTTTTACCTGGGTTTTCAG  615

Qy    553  gatgttggtgcttggtgttgccttggtgtctgtgagagtatacttcaaaaagtgcccattt  612
           ||||| ||| ||| | ||  || || |||| |  |||||||| ||| || || ||||| ||
Db    616  GATGTGGGTCCTTGCATCCCCTCGGTGTCGTCCGTGTGTTCTACAGAAGAAGTGTCCACTC  675

Qy    613  acaagtgaagaaatctggcatggtgtcttcccagacacggtacccatg--gactccgtccctg  669
           ||| ||| ||| |||| | || ||||| | ||||||||  |||||||||    |||  |||||
Db    676  ACGGTTCGAAATCTAGCCCAGTTTCCTGACACCATTACTGGGGCGTAACCTCTTCCCTG  735

Qy    670  gtggaggttagaggtcttgttgtcaacaatttcaaggaggaagatcccaaggatgtac  729
           ||||| ||| || | || | |||||||||||| ||||| |||||| | || |||||
Db    736  GTGGAAAGTTCGAGCGGCTCCGTGTGTCAACAACTCAGAAGAAGAAGATGTGCCAAAAATGTAC  795

Qy    730  tgcagtacagaaggcgaatggcttgtgtaccccattggccaagtgtcctgcaatgctggctat  789
           || | ||| ||||| ||||||| ||| ||||||||||| || || | || |   | ||| |
Db    796  TGTGGGGCAGATGGTGAAAGGCTGGTGTACCCCATTGGCAACCGCTGTGGGGAC  855

Qy    790  gaagaaaagaggtttatgtgccaa  813
           || || |||| | |||| ||||||
Db    856  GAGGAGCAGAAATGGTGAAATGCCAA  879
```

```
RESULT 15
LOCUS       S57168      4242 bp   mRNA            ROD    28-JUN-1993
DEFINITION  Sek=Eph-related receptor protein tyrosine kinase [mice, mRNA, 4242
            ntl.
ACCESSION   S57168
VERSION     S57168.1   GI:298687
KEYWORDS    .
SOURCE      Mus sp.
ORGANISM    Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4242)
AUTHORS     Gilardi-Hebenstreit,P., Nieto,M.A., Frain,M., Mattei,M.G.,
            Chestier,A., Wilkinson,D.G. and Charnay,P.
TITLE       An Eph-related receptor protein tyrosine kinase gene segmentally
            expressed in the developing mouse hindbrain
JOURNAL     Oncogene 8 (4), 1103 (1993)
MEDLINE     93205393
REMARK      GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 127779] from the original journal article.
            This sequence comes from Fig. 5.
            *ERRATUM* Gilardi-Hebenstreit et al., Oncogene (1992)7, 2499-2507.
COMMENT
FEATURES             Location/Qualifiers
     source          1..4242
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            58..3018
                     /gene="Sek"
                     /note="Eph-related receptor protein tyrosine kinase"
     CDS             58..3018
                     /gene="Sek"
```

```
/note="Eph-related receptor protein tyrosine kinase;
*ERRATUM* Gilardi-Hebenstreit et al., Oncogene (1992) 7,
2499-2507; This sequence comes from Fig. 5"
/codon_start=1
/protein_id="AAB25836.1"
/db_xref="GI:298688"
/translation="MAGIFIFILFSFLFGICDAVTGSRVPANEVTLLDRSVQGELG
WIASPLEGGWEEVSIMDEKNIPIRTYQVCNVMEASQNNWLRTDWIIREGAQRVYIEIK
FTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRESQFGKTDTIAADESFTQVDIG
DRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVYYKKCPLIVRNLAQFPDTI
TGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEEQNGECQAC
KIGYYKALSTDASCAKCPPHSYSVWEGATSCTCDRGFFRADNDAASMPCTRPPSAPLN
LISNVNETSVNLEWSSPQNTGGRQDISYNVVCKKCGAGDPSKRCRPCGGSVHYTPQONG
LKTTRVSIIDLLAHTNYTFEIWAVNGVSKYNPSPDGSVSVTVTNQAAPSSIALVQAK
EVTRISVALAMLEPDRPNGVILEYEVKYYEKDQNERSYRIVRTAARNTDIGLNPLIS
YVFHVRARTAAGYGDFSEPLEVTTNTVPSRIIGDANSTVLLVSVSGSVVLVILLIAA
VFISRRSRKYSKAKQEADEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCIKIE
KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKQRRDFLSEASIMGQFDHPNI
IHLEGVVTKCKPVMIITEYMENGSLDAFLRKNDGRFTVIQLVGMLRGIGSGMKYLSDM
SYVHRDLAARNILVSNLCVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
RKFTSASDVWSYGIVMWEVMSYGERYPWDMSNQDVIRAIEEGYRLPPPMDCPIALHQL
MLDCWQKERSDRPKFGQIVNNLDKLIRNPNSLKRTSESESRPNTALLDPSSPEFSAVV
SVGDWLQAIKMDRYKDNFTAAGYTTLEAVHMSQDDLARIGITAIHQNKILSSVQAM
RTQMQOMHGRMVPV"

Query Match      43.2%;  Score 351.2;  DB 12;  Length 4242;
Best Local Similarity  66.5%;  Pred. No. 6.7e-98;
Matches 535;  Conservative 0;  Mismatches 263;  Indels 6;  Gaps 2;

BASE COUNT   1155 a   1034 c   1052 g   1001 t
ORIGIN

Qy   16 tccatcctctcctccagctgctcggaacttcgggaactgattccgag 75
          |||||||  |  | |||| |  ||||| || ||||  |  || | |
Db   76 TTCATCCTCTTTTCGTTCTCTTTTGGAATTCGACCTGTCACCGGTATAC 135

Qy   76 ccttccaatgaagtcaatctactggattcaaaaacaattcaagggaactgtgggctggatc 135
          ||||| || |||| ||| | | |||||| || || | |||||| | |||| ||||||||
Db  136 CCCGCGAATGAAGTTACTTTATTGGATTCCAGATTCGTTCAGGGAGAGCTTGGGTGGATA 135

Qy  136 tcttatccat--cacatgggtgggaagagatcagtggttgtgtggatgaacattacacacc 192
          |          | ||||| ||||| || ||| |||| | |||| |||| |||| ||||| |
Db  196 GCAAGCCCTCTCGGAAGAGGGTGGGAGGGAAGTGAGCAATTATGGATGAGAAAAATACACCG 255

Qy  193 atcaggacttaccagtgtgtgcaatgtcaatgtgaccacagtcaaaacaattggctgagaaca 252
          ||| | |||||||||||||  ||| ||||| | || |||||||| |  ||||||||||| |||
Db  256 ATCCGAACCTACCAGTGTGCAAACGTGATCGAAGCCCAGCCAGACAACTGGCTGCGGAACT 315

Qy  253 aactgggtcccagggaactcagctcagaagatttatgtgggagctcaactctacga 312
          |||||| |  ||| |  | |||| | || |||||||| |||| |||| ||||||||
Db  316 GACTGGATCACCCGCGAAGGGCGCAGAGGGTGTACATTGAAATTAAGTTCACTCTAAGG 375

Qy  313 gactgcaatagacattccattggtttaggaacttcaagaggagacattcaacctgtactac 372
          |||||||| |  |  ||| | ||| | |   | || ||| || |||| ||||||||||||||
Db  376 GACTGCAAACTACTCTTCCGGGGGTCATTGCAAGGAGACGTTTAACCTCTACTAC 435

Qy  373 atggagtctgatgatcatggggtgaaattccgagggcatcacagttacaagattgac 432
          | |||||| || | | | || | || | |||| | |  ||||| |||||| |||||
Db  436 TATGAGTCGGACAACGACACGAGCGGCTTCAACGAGAGCCAGTTTGGCAAGACTGAC 495

Qy  433 accaatgcaagtcgatgaaagtttcactcaaatggatctttgggaacgtatctgaagctc 492
          |||| | |||| || | |||||| ||| ||  || |  | ||| |||| ||| ||||||||
Db  496 ACCATCCCAGCTGATGAGAGCTTCACCCAGGTGACAATGGTCGACCGAATCATGAAAACTC 555

Qy  493 aacactgagattagagaagtaggtcctgtcaacaagaagggattttattttggcattcaa 552
          |||| |||| || ||| || ||| ||  | | |||||  ||||||| | | |||| ||||
Db  556 AACACGAGATCCGGGATGTAGGGCCACTGAGCCAAGAAAAAGGGTTTTACCTGCTTTCAG 615

Qy  553 gatgttggtcgttgttgttcttggtgtctgtgagagtatactccaaaatgcccattt 612
          |||| | | ||  | | ||  | ||||||| | | | | ||||| || |||  |  | |
Db  616 GATGTGGGTCGGTGCTTGCATCGCCCTGGTGTCTGTCCGGTTGTTCTACAAGAAGTGTCCACTC 675

Qy  613 acagtgaagaatctgtgctatgttccaagacaacggtaccccagtg--gactccagtccctg 669
          || ||  | |||||| |   ||| || || | |  | || |  ||||   ||||| || ||||
Db  676 ACCGTTCGAAAATCTAGCCCAGAGTTTCCTGACACCATTACTGGGGCTGATACCTTCTCCCTG 735

Qy  670 gttgaggttagagggttcttgtcaacaattctaagaggaagaatcctccaaggaagatgtac 729
          ||||| || || || || | ||  || | || | || |||| ||| || ||||| |||||||
Db  736 GTGGAAGTTCGAGCGCTCCGTGTGTCAACAACTCAGAAGAGAAGGAATGCCAAAAATGTAC 795

Qy  730 tgcagtacagaaggcgaatggcttgtacccattggcaagtgtcctgcaagtgtggctat 789
          || |  ||||| || ||||| |||||||||||||||||| ||| | ||||  ||||| |
Db  796 TGTGGGGCAGATGGTGAATGGCTGGTACCCATTGGCAACTGCCTATGCAAACGCTGGGCAC 855

Qy  790 gaagaaagaggtttatgtgccaa 813
          || ||| || || |||||||||
Db  856 GAGGAGCAGAATGGTGAATGCCAA 879
```

Search completed: May 15, 2000, 11:42:02
Job time: 18549 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 11:20:11 ; Search time 1358.94 Seconds
                   (without alignments)
                   259.489 Million cell updates/sec

Title: US-09-104-340-6
Perfect score: 87
Sequence: 1 atggattgtcagcttccat...........ttccgcagccttccaatgaa 87

Scoring table: IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   EST:*

1: em_est1:*	45: gb_est26:*	73: gb_est54:*
2: em_est2:*	46: gb_est27:*	74: gb_est44:*
3: em_est3:*	47: gb_est28:*	75: em_est31:*
4: em_est4:*	48: gb_est29:*	76: em_est32:*
5: em_est5:*	49: gb_est30:*	77: em_est33:*
6: em_est6:*	50: gb_est31:*	78: em_est34:*
7: em_est7:*	51: gb_est32:*	79: em_gss2:*
8: em_est8:*	52: em_est25:*	80: gb_est46:*
9: em_est9:*	53: gb_est29:*	81: gb_est47:*
10: em_est10:*	54: em_est22:*	82: gb_est1:*
11: em_est11:*	55: gb_est33:*	83: gb_est2:*
12: em_est12:*	56: em_est25:*	84: gb_est3:*
13: em_est13:*	57: em_est25:*	85: gb_est4:*
14: em_est14:*	58: em_est26:*	86: em_gss1:*
15: em_est15:*	59: gb_est33:*	87: em_gss2:*
16: em_est16:*	60: gb_est34:*	88: em_gss3:*
17: em_est17:*	61: gb_est35:*	89: em_gss4:*
18: em_est18:*	62: gb_est36:*	90: gb_gss5:*
19: em_est19:*	63: gb_est37:*	91: gb_gss6:*
20: em_est1:*	64: gb_est38:*	92: gb_gss7:*
21: gb_est2:*	65: em_est27:*	93: gb_gss8:*
22: gb_est3:*	66: em_est28:*	94: gb_gss9:*
23: gb_est4:*	67: em_est29:*	95: em_gss5:*
24: gb_est5:*	68: em_est30:*	96: em_gss6:*
25: gb_est6:*	69: gb_est39:*	97: em_gss7:*
26: gb_est7:*	70: gb_est40:*	98: em_gss8:*
27: gb_est8:*	71: gb_est41:*	99: em_gss9:*
28: gb_est9:*	72: gb_est42:*	100: em_gss10:*
29: gb_est10:*		101: em_gss11:*
30: gb_est11:*		102: gb_gss10:*
31: gb_est12:*		103: gb_gss11:*
32: gb_est13:*		104: gb_gss12:*
33: gb_est14:*		105: gb_gss12:*
34: gb_est15:*		106: gb_gss13:*
35: gb_est16:*		107: gb_gss14:*
36: gb_est17:*		108: gb_gss15:*
37: gb_est18:*		109: gb_gss16:*
38: gb_est19:*		
39: gb_est20:*		
40: gb_est21:*		
41: gb_est22:*		
42: gb_est23:*		
43: gb_est24:*		
44: gb_est25:*		

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result   Query   SUMMARIES
          %

RESULT 1

```
LOCUS         H38363
DEFINITION    H38363  435 bp  mRNA  EST  16-AUG-1995
              yp50g04.r1 Soares retina N2b4HR Homo sapiens cDNA clone
              IMAGE:190902 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
              RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION     H38363
VERSION       H38363.1  GI:907862
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 435)
AUTHORS       Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
              Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
              Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
              Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P., and
              Wilson,R.
TITLE         The WashU-Merck EST Project
JOURNAL       Unpublished (1995)
```

	No.	Score	Match	Length	DB	ID	Description
	1	82.4	94.7	435	23	H38363	H38363 yp50g04.r1
c	2	30.4	34.9	427	39	AA865380	AA865380 oh49f01.s
c	3	28.8	33.1	225	38	AA752361	AA752361 96GS0872
c	4	28.2	32.9	407	91	AQ850374	AQ850374 LMAJFv11_1
c	5	28.2	32.4	392	43	AU033295	AU033295
c	6	28	32.2	322	50	AV044909	AV044909
c	7	28	32.2	473	74	AV393461	AV393461
c	8	28	32.2	518	74	AW230735	AW230735
c	9	27.8	32.0	307	60	AV153845	AV153845 uo67c05.y
c	10	27.6	31.7	291	83	AF116297	AF116297 Armillari
c	11	27.6	31.7	402	40	AA952881	AA952881 on76f10.s
c	12	27.6	31.7	419	31	AA282730	AA282730 zs89h08.r
c	13	27.6	31.7	484	63	AI992021	AI992021 w199d07.x
c	14	27.6	31.7	549	82	AI865632	AI865632 ws44b02.x
c	15	27.6	31.7	230	38	AA789282	AA789282 aj28c02.s
c	16	27.2	31.3	349	20	T10179	T10179 seq1347 b4H
c	17	27.2	31.3	377	39	AA903650	AA903650 ok63f04.s
c	18	27.2	31.3	388	26	W33901	W33901 mc56f02.r1
c	19	27.2	31.3	401	92	AQ910535	AQ910535 GSSTc0577
c	20	27.2	31.3	455	22	R37827	R37827 yh97e07.s1
c	21	27.2	31.3	468	39	AA825079	AA825079 L30-102T3
c	22	27.2	31.3	471	37	AA709190	AA709190 a120e00.s
c	23	27.2	31.3	479	43	AI187362	AI187362 qf29b07.x
c	24	27.2	31.3	486	33	AA426374	AA426374 zv54h01.s
c	25	27.2	31.3	544	92	AA412719	AA412719 zt67d03.s
c	26	27.2	31.3	554	92	AQ920797	AQ920797 RPCI-23-2
c	27	27.2	31.3	597	80	AW320797	AW320797 uo23f02.y
c	28	27.2	31.3	600	43	AI207994	AI207994 gg51d10.x
c	29	27.2	31.3	630	61	AI822984	AI822984 L30-83T3
c	30	27.2	31.3	674	79	AW266010	AW266010 L30-267IT
c	31	27.2	31.3	701	43	AI208021	AI208021 gg60g06.x
c	32	27.2	31.3	708	43	AI205604	AI205604 gg30f11.x
c	33	27.2	31.3	708	20	Z43437	Z43437 HSC1CA061 n
c	34	26.8	30.8	270	20	AI331075	AI331075 fa95h11.y
c	35	26.8	30.8	462	45	AL134841	AL134841 DKFZp762C
c	36	26.8	30.8	564	79	AJ273151	AJ273151 AJ273151
c	37	26.8	30.8	613	74	FR000603	FR000603 F.rubripes
c	38	26.8	30.8	619	82	Z89847	Z89847 F.rubripes
c	39	26.6	30.6	619	70	AW144998	AW144998 rs44d10.y
c	40	26.6	30.6	280	21	R05766	R05766 ye88g05.r1
c	41	26.6	30.6	315	93	AQ988434	AQ988434 RPCI-23-3
c	42	26.6	30.6	395	47	AI518228	AI518228 LD37682.5
c	43	26.6	30.6	458	47	AI512933	AI512933 LD45068.5
c	44	26.6	30.6	490	47	AI512463	AI512463 LD44345.5
c	45	26.6	30.6	545	47	AI546426	AI546426 LD47994.5

ALIGNMENTS

```
COMMENT       On May 5, 1995 this sequence version replaced gi:797807.
              Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Insert Size: 2663
              High quality sequence stops: 276
              Source: IMAGE Consortium, LLNL
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 2663   Std Error: 0.00
              Seq primer: M13Rev
              High quality sequence stop: 276.
FEATURES             Location/Qualifiers
     source          1..435
                     /organism="Homo sapiens"
                     /db_xref="GDB:3847311"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:190902"
                     /clone_lib="Soares retina N2b4HR"
                     /sex="male"
                     /tissue_type="retina"
                     /dev_stage="55 year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGAGGCGGCGCGGGCGCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). The retinas were obtained from a 55 year old
                     Caucasian and total cellular poly(A)+ RNA was kindly
                     provided by Roderick R. McInnes M.D. Ph.D. from the
                     University of Toronto. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      109 a    108 c    103 g    109 t    6 others
ORIGIN
```

```
Query Match     94.7%;  Score 82.4;  DB 23;  Length 435;
Best Local Similarity   97.68%;  Pred. No. 8.4e-14;
Matches   83;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy    3 ggattgtcaagtctccatcctcctccttcagctgctcgttcctcgacagcttcgggga   62
        |||||||||| ||||||||||| |||||||||||||||||||||||| ||||||||||||
Db   69 GCATTGTCAGCTCTCCAATCCCTCCTCCTTCAGCTGCTCGTTCCTCGACACAGCTTCGGGGA 128

Qy   63 actgattcccgcagcctccaatgaa   87
        ||| |||||||||||||||||||||||
Db  129 ACTNATTCCGCAGCCTTCCAATGAA  153
```

RESULT 2

```
LOCUS         AA865380   427 bp   mRNA   EST   12-MAR-1998
DEFINITION    oh49f01.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1469977 3'
              similar to gb:M17887 60S ACIDIC RIBOSOMAL PROTEIN P2 (HUMAN);, mRNA
              sequence.
ACCESSION     AA865380
VERSION       AA865380.1   GI:2957656
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 427)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
```

JOURNAL   Unpublished (1997)
COMMENT   On Jan 14, 1998 this sequence version replaced gi:1798438.
          Contact: Robert Strausberg
          Tel: (301) 496-1550
          Email: Robert_Strausberg@nih.gov
          Tissue Procurement:Christopher A. Moskaluk, M.D., Ph.D., Michael
          Emmert-Buck, M.D., Ph.D.
          cDNA Library Preparation: M. Bento Soares, Ph.D.
          cDNA Library Arrayed by: Greg Lennon, Ph.D.
          DNA Sequencing by: Washington University Genome Sequencing Center
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 373.

FEATURES   Location/Qualifiers
     source   1..427
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1469977"
                     /clone_lib="NCI_CGAP_GC4"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo."

BASE COUNT   89 a   114 c   115 g   109 t
ORIGIN

RESULT   3
AA75235l/c
LOCUS   AA752361   225 bp   mRNA   EST   20-JAN-1998
DEFINITION   96GS0872 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
             cDNA clone 96GS0872, mRNA sequence.
ACCESSION   AA752361
VERSION   AA752361.1   GI:2799372
KEYWORDS   EST.
SOURCE   Oryza sativa
  ORGANISM   Oryza sativa
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
             Poaceae; Oryza.
REFERENCE   1 (bases 1 to 225)
  AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
             Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
             Lee,M.C. and Eun,M.Y.
  TITLE   Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
  JOURNAL   Unpublished (1998)
COMMENT   On May 8, 1995 this sequence version replaced gi:801205.
             Contact: Eun M.Y.
             Department of Cytogenetics
             National Inst. of Agri. Sci. and Tech, RDA
             Suwon, Kyunggido, Korea

Query Match   34.9%; Score 30.4; DB 39; Length 427;
Best Local Similarity 67.2%; Pred. No. 33;
Matches   43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy   13 ctctccatcctccctccttctcagctgtcctctgttccgacagcttcggg 72
           |||| ||||||||| |||||||| ||||||| ||||| || ||||||
Db   82 CTCTTCAGACTCCTCCTTCTTCTCARCTTTCTCTTCCCTCTGCACAGGGGCCGAGAACCAG 141

Qy   73 cagc 76
           ||||
Db   142 CAGC 145

RESULT   4
AQ850374
LOCUS   AQ850374   407 bp   DNA   GSS   15-OCT-1999
DEFINITION   LMAJFV1_lm30b08.y1 Leishmania major FV1 random genomic library
             Leishmania major genomic clone LMAJFV1_lm30b08 5', genomic
             sequence.
ACCESSION   AQ850374
VERSION   AQ850374.1   GI:6055022
KEYWORDS   GSS.
SOURCE   Leishmania major
  ORGANISM   Leishmania major
             Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
             Leishmania.
REFERENCE   1 (bases 1 to 407)
  AUTHORS   Akopyants,N.S., Beverley,S.M., Clifton,S., Marra,M., Hillier,L.,
             Chinwalla,A., Pape,D., Martin,J., Wylie,T., Blistain,A.,
             Schmitt,A., Person,B., Thelsing,B., Ritter,E., Ronko,I.,
             Bennett,J., Cole,R., Underwood,K., Cardenas,M., Gibbons,M.,
             Harvey,N., McCann,R., Tsagareishvili,R., Williams,T., Jackson,Y.,
             Bowers,Y., Swaller,T., Waterston,R. and Wilson,R.
             WashU Leishmania Project
  TITLE   WashU Leishmania Project
  JOURNAL   Unpublished (1999)
COMMENT   On Mar 23, 1999 this sequence version replaced gi:3323768.
             Other GSSs: lm30b08.x1
             Contact: Akopyants, NS / Beverley, SM
             WashU Leishmania Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             Library construction: Natalia S. Akopyants, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             If using this information please cite:
             N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
             Friedlin strain V1 genome by shotgun sequencing' and the Washington
             University Genome Sequencing Center for information on obtaining

---

Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea, 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse primer.

FEATURES   Location/Qualifiers
     source   1..225
                     /organism="Oryza sativa"
                     /cultivar="Milyang23"
                     /db_xref="taxon:4530"
                     /clone="96GS0872"
                     /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
                     /tissue_type="Immature Seed"
                     /dev_stage="5 days after pollination"
                     /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                     XhoI; Directional cDNA library inserted into lambda ZAPII
                     vector at 5'end with EcoRI and 3' end with Xho I site."

BASE COUNT   53 a   69 c   45 g   58 t
ORIGIN

Query Match   33.1%; Score 28.8; DB 38; Length 225;
Best Local Similarity 65.6%; Pred. No. 77;
Matches   42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1 atggatgtgtcagctctccatctctccttcagctgtctgttccgacagcttcggg 60
           |||||| |||||| ||||| ||||||| |||| | ||||| |||||||| ||
Db   157 ATGGATTTTCCAGTCTCCAGTANCCACCCTTCACCGTCTCACGTTCTCGCTAGAGGCGGA 98

Qy   61 gaac 64
           || ||
Db   97 GGAC 94

clone material please contact: Natalia S. Akopyants Ph.D.
(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.
(beverleye@borcim.wustl.edu)
Seq primer: -40RP from Gibco
Class: shotgun
High quality sequence stop: 365.

```
FEATURES             Location/Qualifiers
     source          1..407
                     /organism="Leishmania major"
                     /strain="Friedlin strain V1"
                     /db_xref="taxon:5664"
                     /clone="LMAJFV1_lm30b08"
                     /clone_lib="Leishmania major FV1 random genomic library"
                     /lab_host="TOP10 (Invitrogen)"
                     /note="Vector: pZero-2 (Invitrogen); Site_1: EcoRV;
                     Genomic DNA was isolated from stationary phase cells. For
                     this library, DNA was sheared to give a tight size
                     distribution of 1-1.5kb fragments, blunt-ended with T4 DNA
                     polymerase, dephosphorylated with Shrimp Alkaline
                     Phosphatase and ligated into pZero-2 vector's EcoRV site."
BASE COUNT       62 a      148 c     107 g      90 t
ORIGIN
```

```
RESULT 5
Query Match           32.9%;    Score 28.6;  DB 91;  Length 407;
Best Local Similarity 61.3%;    Pred. No. 1.1e+02;
Matches   46;  Conservative   0;  Mismatches  29;  Indels  0;  Gaps  0;

Qy    11 agctctccaatcctcctccttccagctcgtctgttcctcgacagcttcggggaactgattc 70
          ||| |  |||||| | ||   |  | ||| |||||||||||||| ||| |  |  |  || ||
Db   185 AGCGCCTGTGCTCCTCCTTCAGAGCTTCTATGCCGGCGGGCAATAACCTGTAGTTGATGT 244

Qy    71 cgcagccttcaatg 85
          ||| || |||| ||
Db   245 CGCTGCCGTCCACTG 259
```

```
LOCUS       AU033295      392 bp    mRNA            EST       20-OCT-1998
DEFINITION  AU033295 Rice shoot Oryza sativa cDNA clone S6421_1A, mRNA
            sequence.
ACCESSION   AU033295
VERSION     AU033295.1  GI:3769268
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Oryza.
REFERENCE   1  (bases 1 to 392)
  AUTHORS   Yamamoto,K. and Sasaki,T.
  TITLE     Rice cDNA from etiolated shoot
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 17, 1998 this sequence version replaced gi:1900162.
            Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program
            2-1-2 Kannondai,Tsukuba
            Ibaraki,
            Japan 305
            Tel: 0298-38-7441
            Fax: 0298-38-7468
            Email: tsasaki@abr.affrc.go.jp
            PROJECT = RGP;
FEATURES             Location/Qualifiers
     source          1..392
                     /organism="Oryza sativa"
                     /strain="Nipponbare, sub_species Japonica"
                     /db_xref="taxon:4530"
                     /clone="S6421_1A"
                     /clone_lib="Rice shoot"
                     /note="Etiolated shoot (8 days old)"
BASE COUNT       85 a      80 c     125 g      95 t       7 others
ORIGIN

Query Match           32.4%;    Score 28.2;  DB 43;  Length 392;
Best Local Similarity 58.5%;    Pred. No. 1.3e+02;
Matches   48;  Conservative   0;  Mismatches  34;  Indels  0;  Gaps  0;

Qy     1 atggatgtcagctctccaatcctccttccttccagctcgtctcgttcctcgacagcttcggg 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   109 ATAGCTTCGGCCACCTTCTTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCC 50

Qy    61 gaactgattccgcagccttcca 82
          |||   |||| |  || ||
Db    49 GAGGAAGCTCCCGGGGCTAATTCA 28
```

```
RESULT 6
LOCUS       AV044909      322 bp    mRNA            EST       23-NOV-1999
DEFINITION  AV044909 Mus musculus adult C57BL/6J testis Mus musculus cDNA clone
            1700040E03, mRNA sequence.
ACCESSION   AV044909
VERSION     AV044909.2  GI:4864574
KEYWORDS    EST.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            house mouse.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 322)
  AUTHORS   Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
            Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
            Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
            Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
            Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
            Sugahara,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y.,
            Tominaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
            Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs
  JOURNAL   Unpublished (1999)
COMMENT     On Jun 5, 1998 this sequence version replaced gi:3187926.
            Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-res@rtc.riken.go.jp
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length cDNA
            (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1..322
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="1700040E03"
                     /clone_lib="Mus musculus adult C57BL/6J testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
BASE COUNT       79 a      96 c      62 g      85 t
ORIGIN

Query Match           32.2%;    Score 28;  DB 50;  Length 322;
Best Local Similarity 63.2%;    Pred. No. 1.4e+02;
Matches   43;  Conservative   0;  Mismatches  25;  Indels  0;  Gaps  0;
```

```
Qy    8 gtcagctctccatcctcctcctcctctagctgtctcgttcttctcgacagcttcgggggaactga   67
        | ||| |||   ||||||||| |||||| ||||||| ||| ||| || ||||| |||| |||
Db  172 GCCACCTTCCTCCACCTTCTCAGCTCTCCTCAAATTCTCCCTCTTCCTCAGCTGTGGGGTCCTGG  113

Qy   68 ttccgcgag  75
        ||| |||
Db  112 TACTGCTG  105
```

RESULT 7

AV393461/c

LOCUS   AV393461   473 bp   mRNA   EST   21-JAN-2000
DEFINITION   AV393461 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
             cDNA clone CL07e07_r, mRNA sequence.
ACCESSION   AV393461
VERSION   AV393461.1 GI:6547677
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii.
  ORGANISM   Chlamydomonas reinhardtii
             Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
             Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 473)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S., Fukuzawa,H. and Tabata,S.
  TITLE   A Large Scale Structural Analysis of cDNAs in a Unicellular Green
             Alga, Chlamydomonas reinhardtii. I. Generation of 3433
             Non-redundant Expressed Sequence Tags
  JOURNAL   DNA Res. 6, 369-373 (1999)
COMMENT   On Jul 9, 1999 this sequence version replaced gi:5434219.
             Contact: Yasukazu Nakamura
             The First Laboratory for Plant Gene Research
             Kazusa DNA Research Institute
             Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
             Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES   Location/Qualifiers
     source   1..473
                     /organism="Chlamydomonas reinhardtii"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="CL07e07_r"
                     /clone_lib="Chlamydomonas reinhardtii C9"
                     /dev_stage="photoautotrophic growth"
                     /note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT   83 a   151 c   151 g   88 t
ORIGIN

```
Qy    2 tggattgtcagctctccatcctcctcctcctctagctgtctcgttcttctcgacagcttcgggg   61
        ||||| |||| |||   ||||||||| |||||| ||||||| ||| ||| || ||||| ||||
Db  264 TGGATGGTCAGCTCTCCCACCTTCTCTCCTTGGTGAGCTGAAGCCGCGACAGACACAGCCGCGG  205
```

RESULT 8

AW230735/c

LOCUS   AW230735   518 bp   mRNA   EST   10-DEC-1999
DEFINITION   uo67c05.y1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:2647592 5'
             similar to gb:D21089 DNA-REPAIR PROTEIN COMPLEMENTING XP-C CELLS
             (HUMAN);, mRNA sequence.
ACCESSION   AW230735
VERSION   AW230735.1 GI:6560031
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 518)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
```

MGI:1028044
Seq primer: -40RP from Gibco
High quality sequence stop: 442.

FEATURES   Location/Qualifiers
     source   1..518
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2647592"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT   169 a   116 c   163 g   68 t   2 others
ORIGIN

Query Match   32.2%;   Score 28; DB 74;   Length 518;
Best Local Similarity 66.7%;   Pred. No. 1.6e+02;
Matches   40; Conservative   0; Mismatches 20; Indels   0; Gaps   0;

RESULT 9

AV153845/c

LOCUS   AV153845   307 bp   mRNA   EST   08-JUL-1999
DEFINITION   AV153845 Mus musculus hippocampus C57BL/6J adult Mus musculus cDNA
             clone 2900058C14, mRNA sequence.
ACCESSION   AV153845
VERSION   AV153845.1 GI:5359051
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 307)
  AUTHORS   Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
             Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
             Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
             Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
             Owa,C., Sato,K., Shibata,Y., Suzuki,H., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
             Sugahara,Y., Suzuki,H., Suzuki,Y., Tateno,M., Tomaru,Y.,
             Tominaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
             Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE   RIKEN Mouse ESTs
  JOURNAL   Unpublished (1999)
COMMENT   On May 18, 1998 this sequence version replaced gi:3137133.
             Contact: Chie Owa
             Genome Science Laboratory
             RIKEN
             3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
             Tel: 81-298-36-9145
             Fax: 81-298-36-9098

Query Match   32.2%;   Score 28; DB 74;   Length 473;
Best Local Similarity 66.7%;   Pred. No. 1.7e+02;
Matches   40; Conservative   0; Mismatches 20; Indels   0; Gaps   0;

```
FEATURES
     source          Location/Qualifiers
                     1..307
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2900058C14"
                     /clone_lib="Mus musculus hippocampus C57BL/6J adult"
                     /sex="male"
                     /tissue_type="hippocampus"
                     /dev_stage="adult"
BASE COUNT       82 a       75 c       80 g       70 t
ORIGIN
```

```
Qy    7 tgtcagctctcatcctcctcctcttctcagctgtctctgttctcgacagcttcggggaactg 66
        |||||||| ||| |||||||||||||||||| ||| |||| |||||||||||| |||||
Db   88 TGCCAGCTCTCCAATCCTCAGCGTTCTTCGACTGCTCTCCTTGAAAAGTTGTGAGGAAAC 29

Qy   67 att 69
        |||
Db   28 ATT 26
```

```
RESULT 10
LOCUS       AF116297     291 bp     DNA                 GSS     10-MAY-1999
DEFINITION  Armillaria gallica strain C33r, genomic survey sequence.
ACCESSION   AF116297
VERSION     AF116297.1  GI:4761819
KEYWORDS    GSS.
SOURCE      Armillaria gallica
  ORGANISM  Armillaria gallica
            Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
            Tricholomataceae; Armillaria.
REFERENCE   1  (bases 1 to 291)
  AUTHORS   Anderson,J.B. and Hodnett,B.
  TITLE     Random nuclear DNA sequences of Armillaria gallica from a northern
            Michigan hardwood forest
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 291)
  AUTHORS   Anderson,J.B. and Hodnett,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-DEC-1998) Biology, University of Toronto at
            Mississauga, 3359 Mississauga Road, Mississauga, Ontario L5L 1C6,
            Canada
FEATURES
     source          Location/Qualifiers
                     1..291
                     /organism="Armillaria gallica"
                     /strain="C33r"
                     /db_xref="taxon:47427"
BASE COUNT       30 a       94 c       72 g       93 t       2 others
ORIGIN
```

Query Match   31.7%;   Score 27.6; DB 83; Length 291;
Best Local Similarity 63.6%;   Pred. No. 1.8e+02;
Matches   42;   Conservative 0;   Mismatches 24;   Indels 0; Gaps 0;

```
Qy   13 cctccatccctcctccagctggttctcgacagcttcggggaactgattccg 72
        |||| ||||||||||| ||||||||| ||||||||||| ||  ||| ||
Db  152 CTCTTCAATCCTCTCTCCACTTCTCTTCTTCTTCGTCGTCGTCGTCAAAATTCGGCTCGG 211
```

Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

```
Qy   73 cagcct 78
        || |||
Db  212 GTGCTT 217
```

```
RESULT 11
LOCUS       AA952881     329 bp     mRNA                EST     23-JUN-1998
DEFINITION  on76fl0.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1562635 3', mRNA sequence.
ACCESSION   AA952881
VERSION     AA952881.1  GI:3117028
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   On Jan 17, 1998 this sequence version replaced gi:2044427.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 704    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES
     source          Location/Qualifiers
                     1..329
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1562635"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT       77 a      102 c       82 g       68 t
ORIGIN
```

Query Match   31.7%;   Score 27.6; DB 40; Length 329;
Best Local Similarity 60.8%;   Pred. No. 1.9e+02;
Matches   45;   Conservative 0;   Mismatches 29;   Indels 0; Gaps 0;

```
Qy   14 tctccatccctcctttctcagctgctctcgttctcgacagcttcggggaactgattccgc 73
        ||||| |||||| ||||| ||| ||||||| ||||||||||||| |||||||| | ||
Db  174 TCTCCAGCCTCAGCCTTCCCAGACCCCCACACACAGGCCACCAGCCTCTGGGAAGTACTGGGAC 233
```

```
Qy   74 agccttccaatgaa 87
        ||| ||| |||||
Db  234 AGGCCTCGGAGCAA 247
```

```
RESULT 12
LOCUS       AA282730     402 bp     mRNA                EST     13-AUG-1997
DEFINITION  zs89h08.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:704703 5'
            similar to TR:G1049078 G1049078 SRP30C. ;, mRNA sequence.
ACCESSION   AA282730
```

VERSION   AA282730.1 GI:1925556
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 402)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On May 8, 1995 this sequence version replaced gi:800900.
             Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email:Robert_Strausberg@nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 960   Std Error: 0.00
             Seq primer: -28m13 rev2 ET from Amersham
             High quality sequence stop: 395.
FEATURES   Location/Qualifiers
     source   1..402
                     /organism="Homo sapiens"
                     /db_xref="GDB:5854427"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:704703"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAANTCTGAAGTGGGAGCGGGCCGCCCATTTTTTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT   102 a   82 c   108 g   110 t
ORIGIN

Query Match   31.7%;   Score 27.6; DB 31; Length 402;
Best Local Similarity 60.8%;   Pred. No. 2e+02;
Matches   45; Conservative   0; Mismatches   29; Indels   0; Gaps   0;

Qy   2 tggattgtcagctctccatcctcctctccagctgctctgttctcgacagcttcgggg 61
           |||||| ||| ||||||||||||||||||||||||||||||||| ||||||| ||||
Db   281 TCGACCANTCCCACTCCCATCCTTCGACANCATGAGCANTAAACAGACANTCCCAGCTTCCGC 222

Qy   62 aactgattccgcag 75
           |||||| || |||
Db   221 ATGTGAATCCCTAG 208

RESULT 13
LOCUS   AI869632   419 bp   mRNA   EST   30-AUG-1999
DEFINITION wl99007.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2433037 3',
           mRNA sequence.
ACCESSION AI869632
VERSION   AI869632.1 GI:5543600
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE   National Cancer Institute / National Institute of Neurological
             Disorders and Stroke, Brain Tumor Genome Anatomy Project
             (CGAP/BTGAP), Tumor Gene Index
JOURNAL   Unpublished (1998)
COMMENT   On Jun 22, 1998 this sequence version replaced gi:3247580.
             Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
             Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
             Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html

             Seq primer: -40UP from Gibco
             High quality sequence stop: 412.
FEATURES   Location/Qualifiers
     source   1..419
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2433037"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAANTCTGAAGTGGGAGCGGGCCGCCATNAGGTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT   96 a   131 c   104 g   88 t
ORIGIN

Query Match   31.7%;   Score 27.6; DB 61; Length 419;
Best Local Similarity 60.8%;   Pred. No. 2e+02;
Matches   45; Conservative   0; Mismatches   29; Indels   0; Gaps   0;

Qy   14 tctccaatcctcctccttctcagctgtctcgttctcgacagcttcggggaactgatttccgc 73
           ||||| ||||||||| ||||| |||| || |||| |||||| ||||||| ||
Db   180 TCTCCAGCCTCACGCCTTCCCAGAGCCCCACNACCAGGCCAGCCTCTGGGAAGTACTGGGAC 239

Qy   74 agccttccaatgaa 87
           || ||| || ||
Db   240 AGGCCTCGGAGCAA 253

RESULT 14
LOCUS   AI992021   484 bp   mRNA   EST   08-SEP-1999
DEFINITION ws44b02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2500011 3',
           mRNA sequence.
ACCESSION AI992021
VERSION   AI992021.1 GI:5838926
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 484)
AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute / National Institute of Neurological
           Disorders and Stroke, Brain Tumor Genome Anatomy Project
           (CGAP/BTGAP), Tumor Gene Index
JOURNAL   Unpublished (1998)
COMMENT   On Dec 20, 1995 this sequence version replaced gi:1133839.

```
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
```

```
FEATURES             Location/Qualifiers
     source          1..484
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2500011"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGAGCGGCCGCAATAGGTTTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      116 a    141 c    124 g    103 t
ORIGIN
```

Seq primer: -40UP from Gibco
High quality sequence stop: 471.

```
Qy    14 tctccatcctcctccttcagctgtcctgtgggaacttcggggaactg 73
         |||||||||||||||||||||| |||||| ||||||| ||||||
Db   174 TCTCCACCCTCAGCCTTCCCAGAGCCCCACACCAGGCCGCAGCCTTCGGGAAGTACTGGAGC 233

Qy    74 agccttccaatgaa 87
         |||| || | |||
Db   234 AGGCCTCGGAGCAA 247

Query Match     31.7%;   Score 27.6;   DB 63;   Length 484;
Best Local Similarity   60.8%;   Pred. No. 2.1e+02;
Matches  45;   Conservative  0;   Mismatches  29;   Indels  0;   Gaps  0;
```

```
RESULT 15
FR0019242
LOCUS       FR0019242         549 bp    DNA            GSS       09-DEC-1997
DEFINITION  F.rubripes GSS sequence, clone 027G17aG10, genomic survey sequence.
ACCESSION   AL012136
VERSION     AL012136.1   GI:2677569
KEYWORDS    GSS; genome survey sequence.
SOURCE      Fugu rubripes
  ORGANISM  Fugu rubripes
            Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
            Tetraodontiformes; Tetraodontoidei; Tetraodontidae; Fugu.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   Elgar,G., Clark,M., Meek,S., Warner,S., Umrania,Y.,
            Williams,G., Smith,S., and Brenner,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmp.mrc.ac.uk
COMMENT     Vector: pBluescript II KS
            V_type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
```

```
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Fugu rubripes"
                     /db_xref="taxon:31033"
                     /clone_lib="cosmid 027G17"
                     /clone="027G17aG10"
BASE COUNT      131 a     98 c    140 g    140 t    11 others
ORIGIN
```

```
Qy     9 tcagctccatcctcctccttcagctgtcctgttctcgacaagcttcggggaactg 66
         |||||| ||  |||||   | ||||  || ||  ||| | ||| || || |||
Db    82 TCAGCTGTGCTTCCTCCACCTTCAGCAGTGTCGTGGGTTTTTGCCTGCTTACGGGTCCTG 139

Query Match     31.7%;   Score 27.6;   DB 82;   Length 549;
Best Local Similarity   67.2%;   Pred. No. 2.2e+02;
Matches  39;   Conservative  0;   Mismatches  19;   Indels  0;   Gaps  0;
```

Search completed: May 15, 2000, 11:20:15
Job time: 19138 sec

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 13, 2000, 15:18:09 ; Search time 1681.72 Seconds
                 (without alignments)
                 -160.230 Million cell updates/sec

Title: US-09-104-340-6
Perfect score: 87
Sequence: 1 atggattgtcagctctccat.........ttccgcagccttccaatgaa 87

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 5414606 seqs, -154632823 residues

Total number of hits satisfying chosen parameters: 10829192

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   Pending_Patents_NA:*

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081A_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US081B_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US081C_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US081D_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US082A_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US082B_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US082C_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US083A_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US083B_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US084A_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US084B_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US084C_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US087A_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US087B_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US087C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US088A_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US088B_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US088C_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US088D_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US089A_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US089B_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US089C_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US089D_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US089E_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US089F_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US089G_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US089H_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US089I_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US090A_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US090B_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US090C_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US090D_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US090E_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US091A_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US091B_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US091C_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US092C_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US092D_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US092E_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US093C_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US093D_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US093E_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US093F_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US094A_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US094B_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US094C_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6001A_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6001B_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6001C_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6002A_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6002B_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6002C_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6003A_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6003B_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6004A_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6004B_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6004C_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6008A_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6008B_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6008C_COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US6009A_COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US6009B_COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US6010A_COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US6010B_COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US6011A_COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US6011B_COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US6012A_COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US6012B_COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US6013A_COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US6013B_COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US6014A_COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US6014B_COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US6014C_COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US6016A_COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US6016B_COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US6016C_COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US6017A_COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US6017B_COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US6017C_COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US6018A_COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US6018B_COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US6018C_COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	87	100.0	87	US-09-104-340-6	Sequence 6, Appli
2	87	100.0	813	US-09-104-340-5	Sequence 5, Appli
3	31	35.6	453	US-60-172-375-1200	Sequence 1200, Ap

RESULT 1
US-09-104-340-6
Sequence 6, Application US/09104340
GENERAL INFORMATION:
; APPLICANT: BOYD, Andrew W
; APPLICANT: DOTTORI, Mirella
; APPLICANT: LACKMANN, Martin
; TITLE OF INVENTION: RECEPTOR-LIGAND SYSTEM AND ASSAY
; FILE REFERENCE: boyduq
; CURRENT APPLICATION NUMBER: US/09/104,340
; CURRENT FILING DATE: 1998-06-25
; EARLIER APPLICATION NUMBER: PO7549
; EARLIER APPLICATION FILING DATE: 1997-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(87)
; OTHER INFORMATION: Exon I of HEK gene
US-09-104-340-6

#						
4	28.8	33.1	120625	90	US-60-164-769-24606	Sequence 24606, A
5	28.8	33.1	263456	96	US-60-185-359-1293	Sequence 1293, Ap
6	28.5	33.1	263456	103	US-09-528-237-1293	Sequence 1293, Ap
7	28.5	32.9	310	83	US-08-220-875-1443	Sequence 1443, Ap
8	28.4	32.6	999	57	US-09-514-000-6818	Sequence 6818, Ap
9	28.4	32.6	4266	91	US-60-168-139-2003	Sequence 2003, Ap
10	28.4	32.6	38536	57	US-60-514-000-369	Sequence 369, Ap
11	28.2	32.4	1454	90	US-09-164-320-6260	Sequence 6260, App
12	28.2	32.4	1454	95	US-60-183-791-6260	Sequence 6260, Ap
13	28	32.2	174	103	US-09-531-113-33845	Sequence 33845, A
14	28	32.2	344	86	US-60-144-084-16459	Sequence 16459, Ap
15	28	32.2	391	86	US-60-150-582-260	Sequence 260, Ap
16	28	32.2	419	86	US-60-144-084-17798	Sequence 17798, A
17	27.8	32.0	523	48	US-60-306-609-9784	Sequence 9784, Ap
18	27.8	32.0	523	57	US-09-522-251-9784	Sequence 9784, Ap
19	27.8	32.0	15944	88	US-60-144-084-14736	Sequence 14736, A
20	27.6	31.7	394	87	US-60-146-224-3508	Sequence 3508, Ap
21	27.6	31.7	464	9	US-08-220-691-5570	Sequence 5570, Ap
22	27.6	31.7	464	57	US-08-220-691-5570	Sequence 5570, Ap
23	27.4	31.5	256	57	US-09-540-213-49647	Sequence 49647, A
24	27.4	31.5	283	48	US-09-306-349-22118	Sequence 22118, A
25	27.4	31.5	337	54	US-60-411-999-13357	Sequence 13357, A
26	27.4	31.5	337	89	US-09-105-307A-847	Sequence 847, Ap
27	27.4	31.5	351	87	US-60-162-747-2955	Sequence 2955, Ap
28	27.4	31.5	358	87	US-60-146-224-5923	Sequence 5923, Ap
29	27.4	31.5	559	42	US-60-170-862-21	Sequence 21, Appl
30	27.4	31.5	645	92	US-60-171-432-10092	Sequence 10092, App
31	27.4	31.5	1492	93	US-60-173-383-14401	Sequence 14401, A
32	27.4	31.5	1492	93	US-60-173-464-11278	Sequence 11278, A
33	27.4	31.5	1495	93	US-60-167-217-13895	Sequence 13895, A
34	27.4	31.5	2676	93	US-60-173-383-37298	Sequence 37298, A
35	27.4	31.5	2676	93	US-60-173-464-28401	Sequence 28401, A
36	27.4	31.5	3805	93	US-60-173-383-14400	Sequence 14400, A
37	27.4	31.5	3805	93	US-60-173-464-11277	Sequence 11277, A
38	27.4	31.5	3808	90	US-60-167-217-13894	Sequence 13894, A
39	27.4	31.5	24092	88	US-60-150-584-204	Sequence 204, App
40	27.4	31.5	169516	96	US-60-164-769-24276	Sequence 24276, A
41	27.4	31.5	169519	96	US-60-185-359-997	Sequence 997, App
42	27.2	31.3	169519	103	US-09-528-237-997	Sequence 997, App
43	27.2	31.3	352	89	US-60-160-190-1169	Sequence 1169, Ap
44	27.2	31.3	353	89	US-60-160-190-551	Sequence 651, App
45	27.2	31.3	353	91	US-60-169-841-564	Sequence 564, App

ALIGNMENTS

Query Match 100.0%; Score 87; DB 40; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.7e-16;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1 atggattgtcaagctctccaatcctccatcctgctgctgtctgttcctcggacaagcttcggg 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 atggattgtcaagctctccaatcctccatcctgctgctgtctgttcctcggacaagcttcggg 60

Qy 61 gaactgatccgcagcctccaatgaa 87
       |||||||||||||||||||||||||
Db 61 gaactgatccgcagcctccaatgaa 87
```

RESULT 2
US-09-104-340-5
Sequence 5, Application US/09104340
GENERAL INFORMATION:
; APPLICANT: BOYD, Andrew W
; APPLICANT: DOTTORI, Mirella
; APPLICANT: LACKMANN, Martin
; TITLE OF INVENTION: RECEPTOR-LIGAND SYSTEM AND ASSAY
; FILE REFERENCE: boyduq
; CURRENT APPLICATION NUMBER: US/09/104,340
; CURRENT FILING DATE: 1998-06-25
; EARLIER APPLICATION NUMBER: PO7549
; EARLIER APPLICATION FILING DATE: 1997-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(813)
; OTHER INFORMATION: Exons I, II and III of HEK gene
US-09-104-340-5

Query Match 100.0%; Score 87; DB 40; Length 813;
Best Local Similarity 100.0%; Pred. No. 3.8e-16;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1 atggattgtcaagctctccaatcctccatcctgctgctgtctgttcctcggacaagcttcggg 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 atggattgtcaagctctccaatcctccatcctgctgctgtctgttcctcggacaagcttcggg 60

Qy 61 gaactgatccgcagcctccaatgaa 87
       |||||||||||||||||||||||||
Db 61 gaactgatccgcagcctccaatgaa 87
```

RESULT 3
US-60-172-375-1200/c
Sequence 1200, Application US/60172375
GENERAL INFORMATION:
; APPLICANT: Schuch, Wolfgang
; APPLICANT: Leader, David
; APPLICANT: Doyle, Martin
; APPLICANT: O'Neill, John M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM WHEAT CARYOPSIS
; FILE REFERENCE: PL-0051 P
; CURRENT APPLICATION NUMBER: US/60/172,375
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 11151
; SOFTWARE: PERL Program
; SEQ ID NO 1200
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7020418.1TI

```
;   FEATURE:
;     NAME/KEY: unsure
;     LOCATION: 82, 192, 271, 409
;     OTHER INFORMATION: a, t, c, g, or other
US-60-172-375-1200

Qy     3 ggattgtcagctcagcgccttcca    35.6%;   Score 31; DB 92; Length 453;
         ||||||||  |||| ||||||||
Db   247 GGATACTAAGGCAGCCATGTTCATGCCTCCCAGCTGTTCTGTTATCAACAAACTGNTNGAT 188
Best Local Similarity  35.6%;   Score 31; DB 92; Length 453;
Matches  49; Conservative  0; Pred. No. 13; Mismatches  31; Indels  0; Gaps  0;

Qy    63 actgattccgcagccttcca    82
         |||||||  |||| ||||
Db   187 AGTGANTCATCGACCTGTCA 168
```

RESULT   4
US-60-164-769-24606
; Sequence 24606, Application US/60164769
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containin the Nucleic Acid Sequences
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: CL000144
; CURRENT APPLICATION NUMBER: US/60/164,769
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 27328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24606
; LENGTH: 120625
; TYPE: DNA
; ORGANISM: Drosophila
US-60-164-769-24606

```
Qy    14 tctccatcctcctccagctcgtctgttctcgacagcttcggggaactgattccgc 73
         |||| ||||||||||||| ||| |||||| |||  ||||||||||| ||||||||
Db 100444 tctccctcctcctccttcctgttctcctccggtccggaaacccactccgg 100503
Best Local Similarity  33.1%;   Score 28.8; DB 90; Length 120625;
Matches  42; Conservative  0; Pred. No. 1.3e+02; Mismatches  22; Indels  0; Gaps  0;

Qy    74 agcc 77
         |  |
Db 100504 cggc 100507
```

RESULT   5
US-60-185-359-1293
; Sequence 1293, Application US/60185359
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: CL000284
; CURRENT APPLICATION NUMBER: US/60/185,359
; CURRENT FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 2926
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1293
; LENGTH: 263456
; TYPE: DNA
; ORGANISM: Drosophila
US-60-185-359-1293

```
Query Match  35.6%;   Score 31; DB 92; Length 453;
```

RESULT   6
US-09-528-237-1293
; Sequence 1293, Application US/09528237
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: CL000284
; CURRENT APPLICATION NUMBER: US/09/528,237
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 2926
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1293
; LENGTH: 263456
; TYPE: DNA
; ORGANISM: Drosophila
US-09-528-237-1293

```
Qy    14 tctccatcctcctccagctcgtctgttctcgacagcttcggggaactgattccgc 73
         |||| ||||||||||||| ||| |||||| |||  ||||||||||| ||||||||
Db 234659 tctccctcctcctccttcctgttctcctccggtccggaaacccactccgg 234718
Best Local Similarity  33.1%;   Score 28.8; DB 96; Length 263456;
Matches  42; Conservative  0; Pred. No. 1.5e+02; Mismatches  22; Indels  0; Gaps  0;

Qy    74 agcc 77
         |  |
Db 234719 cggc 234722
```

RESULT   7
US-60-133-875-1443/c
; Sequence 1443, Application US/60133875
; GENERAL INFORMATION:
; APPLICANT: Gairow, Bonnie L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Goralski, Thomas J.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT LIVER
; FILE REFERENCE: PZ-0118 P
; CURRENT APPLICATION NUMBER: US/60/133,875
; CURRENT FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 2916
; SOFTWARE: PERL Program
; SEQ ID NO 1443
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2, 16, 119
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE: -
; OTHER INFORMATION: 70160436OH1
US-60-133-875-1443

```
Query Match  32.9%;   Score 28.6; DB 83; Length 310;
```

Best Local Similarity   67.8%;   Pred. No. 65;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy   14 tctccatcctcctcctccttcagctgtctcgttctcgacagccttcggggaactgaatccg 72
        || | |||||   |||   |||||| ||||||||||||||||||| |||||||||||||||
Db  242 TCAGCCTCCTCACTCATCTCACTCACTGCCGTCTCCATGAGCTCATCGATGAGAACTGAATTCCG 184
```

RESULT 8
US-09-514-000-6818/c
; Sequence 6818, Application US/09514000
; GENERAL INFORMATION:
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
;   FILE REFERENCE: 38-10(15490)B
;   CURRENT APPLICATION NUMBER: US/09/514,000
;   CURRENT FILING DATE: 2000-02-23
;   NUMBER OF SEQ ID NOS: 15034
; SEQ ID NO 6818
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Agrobacterium tumefaciens
Db US-09-514-000-6818

Query Match   60.3%; Score 28.4; DB 57; Length 999;
Best Local Similarity   60.3%; Pred. No. 88;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy    4 gattgtcagctctccatcctcctcctccttcagctgtctcgttctcgacagccttcgggggaa 63
        |||| |||||     |||   ||||||  |||||||||||||||||||| |||||||||
Db  597 GATTACGGGCTCACCTTCCTCGGCCTCGGGCATCGCCATTCCAACTTCATCGCTGGCA 538

Qy   64 ctgattccgcagcctcc 81
        ||||| |||||||||
Db  537 CTGATCCTGATGTATTC 520
```

RESULT 9
US-60-168-139-2003
; Sequence 2003, Application US/60168139
; GENERAL INFORMATION:
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequence And Uses Thereof
;   FILE REFERENCE: 38-21(15490)A
;   CURRENT APPLICATION NUMBER: US/60/168,139
;   CURRENT FILING DATE: 1999-12-01
;   NUMBER OF SEQ ID NOS: 3432
; SEQ ID NO 2003
;   LENGTH: 4266
;   TYPE: DNA
;   ORGANISM: Agrobacterium tumefaciens
US-60-168-139-2003

Query Match   32.6%; Score 28.4; DB 91; Length 4266;
Best Local Similarity   60.3%; Pred. No. 1.1e+02;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy    4 gattgtcagctctccatcctcctcctccttcagctgtctcgttctcgacagccttcgggggaa 63
        |||| |||||     |||   ||||||  ||||||||||||||||||||| |||||
Db  610 gattacgggctcaccttcctcggccatcgccattcccaaacttcatcgctggca 669

Qy   64 ctgattccgcagcctcc 81
        ||||| |||||||||
Db  670 ctgatcctgatgtattc 687
```

RESULT 10
US-09-514-000-369/c
; Sequence 369, Application US/09514000

; GENERAL INFORMATION:
;   APPLICANT: Hinkle, Gregory J.
;   APPLICANT: Slater, Steven C.
;   TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
;   FILE REFERENCE: 38-10(15490)B
;   CURRENT APPLICATION NUMBER: US/09/514,000
;   CURRENT FILING DATE: 2000-02-23
;   NUMBER OF SEQ ID NOS: 15034
; SEQ ID NO 369
;   LENGTH: 38536
;   TYPE: DNA
;   ORGANISM: Agrobacterium tumefaciens
Db US-09-514-000-369

Query Match   32.6%; Score 28.4; DB 57; Length 38536;
Best Local Similarity   60.3%; Pred. No. 1.5e+02;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy     4 gattgtcagctctccatcctcctcctccttcagctgtctcgttctcgacagccttcgggggaa 63
         |||| |||||     |||   ||||||  ||||||||||||||||||||| |||||
Db 14406 GATTACGGGCTCACCTTCCTCGGCCTCGGGCATCGCCATTCCAACTTCATCGCTGGCA 14347

Qy    64 ctgattccgcagcctcc 81
         ||||| |||||||||
Db 14346 CTGATCCTGATGTATTC 14329
```

RESULT 11
US-60-164-320-6260/c
; Sequence 6260, Application US/60164320
; GENERAL INFORMATION:
;   APPLICANT: Bower, Stanley G.
;   APPLICANT: Hinkle, Gregory J.
;   TITLE OF INVENTION: Xanthomonas campestris Genome Sequences and Uses Thereof
;   FILE REFERENCE: 38-10(15804)A
;   CURRENT APPLICATION NUMBER: US/60/164,320
;   CURRENT FILING DATE: 1999-11-10
;   NUMBER OF SEQ ID NOS: 18992
; SEQ ID NO 6260
;   LENGTH: 1454
;   TYPE: DNA
;   ORGANISM: Xanthomonas campestris
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
US-60-164-320-6260

Query Match   32.4%; Score 28.2; DB 90; Length 1454;
Best Local Similarity   64.6%; Pred. No. 1.1e+02;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
Qy     6 ttgtcagctctccatcctcctcctccagctgtctcgttctcgacagccttcggggaact 65
         ||||| ||||||    |||||||||||||||||||||||||||| ||||||| |||
Db  1338 TTCTCCCCTCTCTCCTCCTCCTCTCCTTCACCTCTCCTCCTCCTCCTTTTTTTTTCCT 1279

Qy    66 gattc 70
         |||
Db  1278 CCTTC 1274
```

RESULT 12
US-60-183-791-6260/c
; Sequence 6260, Application US/60183791
; GENERAL INFORMATION:
;   APPLICANT: Bower, Stanley G.
;   APPLICANT: Hinkle, Gregory J.
;   TITLE OF INVENTION: Xanthomonas campestris Genome Sequences and Uses Thereof
;   FILE REFERENCE: 38-10(15804)B
;   CURRENT APPLICATION NUMBER: US/60/183,791
;   CURRENT FILING DATE: 2000-02-22
;   NUMBER OF SEQ ID NOS: 18992
; SEQ ID NO 6260

```
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
US-60-183-791-6260

Query Match        32.4%;  Score 28.2;  DB 95;  Length 1454;
Best Local Similarity   64.6%;  Pred. No. 1.1e+02;
Matches   42;  Conservative  0;  Mismatches  23;  Indels   0;  Gaps   0;

Qy     6 ttgtcaagctccaatccctcctcctttctcaagctgctccggggaact 65
         ||||| ||||||||||||||||||||||||||||||||| ||||||| ||
Db  1338 TTCTTCCCTCCTCCTCCTCCTCCTCTTCACCTCTCCCTTCTTTCTTTTTCCT 1279

RESULT 13
US-09-531-113-33845/c
; Sequence 33845, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 33845
; LENGTH: 174
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700892085H1
US-09-531-113-33845

Qy    66 gattc 70
         |||
Db  1278 CCTTC 1274

Query Match        32.4%;  Score 28;  DB 103;  Length 174;
Best Local Similarity   77.3%;  Pred. No. 89;
Matches   34;  Conservative  0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    10 cagctctccaatccctcctcctttctcaagctgtctcgttctcgacag 53
         ||| | |||||||||||||||||||| ||||||| |||||| | |||||
Db    59 CAGTTTTCCCTCCTCCTTCCCTGCTTTCTGCTTCCACCG 16

RESULT 14
US-60-144-084-16459/c
; Sequence 16459, Application US/60144084
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Buehler, Robert E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Coombs, Brian E.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Nelson, Donald E.
; APPLICANT: Shukla, Hridayabhiramjam
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(15444)B
; CURRENT APPLICATION NUMBER: US/60/144,084
; CURRENT FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 47776
; LENGTH: 344
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3109-014-Q1-K1-A10
US-60-144-084-16459

Query Match        32.4%;  Score 28;  DB 86;  Length 344;
Best Local Similarity   77.3%;  Pred. No. 99;
Matches   34;  Conservative  0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    10 cagctctccaatccctcctcctttctcaagctgtctcgttctcgacag 53
         ||| | |||||||||||||||||||| ||||||| |||||| | |||||
Db   122 CAGTTTTCCCTCCTCCTCCTCCTCCTGCTTCTGCTTTCTGCTTCCACCG 79

RESULT 15
US-60-144-084-17798/c
; Sequence 17798, Application US/60144084
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Buehler, Robert E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Coombs, Brian E.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Nelson, Donald E.
; APPLICANT: Shukla, Hridayabhiramjam
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(15444)B
; CURRENT APPLICATION NUMBER: US/60/144,084
; CURRENT FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 47776
; SEQ ID NO 17798
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3109-053-Q1-K1-D9
US-60-144-084-17798

Query Match        32.4%;  Score 28;  DB 86;  Length 391;
Best Local Similarity   77.3%;  Pred. No. 1e+02;
Matches   34;  Conservative  0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    10 cagctctccaatccctcctcctttctcaagctgctctgttctcgacag 53
         ||| | |||||||||||||||||||| ||||| ||| |||| | |||||
Db    76 CAGTTTTCCCTCCTCCTCCTCCTCCTGCTTCCTGCTTTCTGCTTCCACCG 33

Search completed: May 13, 2000, 15:18:39
Job time: 5212 sec
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 12:05:01 ; Search time 102.62 Seconds
(without alignments)
107.646 Million cell updates/sec

Title: US-09-104-340-6
Perfect score: 87
Sequence: 1 atggattgtcagctctccat...........ttccgcagccttccaatgaa 87

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 226296 seqs, 63486255 residues

Total number of hits satisfying chosen parameters: 452592

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	100.0	3132	1	US-08-167-919A-9	Sequence 9, Appli
2	87	100.0	3132	5	US-08-715-106-9	Sequence 9, Appli
3	27.2	31.3	324	1	US-08-619-542B-34	Sequence 34, Appl
4	26	29.9	1893	7	5438126-1	Patent No. 5438126
5	25.6	29.4	318	3	US-08-619-542B-36	Sequence 36, Appl
6	25.6	29.4	323	4	US-08-619-542B-37	Sequence 37, Appl
7	25.2	29.0	950	3	US-08-901-200A-14	Sequence 14, Appl
8	25.2	29.0	1003	1	US-07-900-364B-13	Sequence 13, Appl
9	25.2	29.0	1003	3	US-07-989-847-11	Sequence 11, Appl
10	25.2	29.0	1723	1	US-07-841-646-28	Sequence 28, Appl
11	25.2	29.0	1723	1	US-07-901-703-10	Sequence 10, Appl
12	25.2	29.0	1723	1	US-08-147-023-28	Sequence 28, Appl
13	25.2	29.0	1723	1	US-08-206-864-3	Sequence 3, Appl
14	25.2	29.0	1723	1	US-08-278-729A-20	Sequence 20, Appl
15	25.2	29.0	1723	1	US-08-480-528A-7	Sequence 7, Appl
16	25.2	29.0	1723	1	US-08-479-666-7	Sequence 7, Appl
17	25.2	29.0	1723	1	US-08-155-343A-20	Sequence 20, Appl
18	25.2	29.0	1723	1	US-08-406-672-20	Sequence 20, Appl
19	25.2	29.0	1723	2	US-08-643-563A-20	Sequence 20, Appl
20	25.2	29.0	1723	2	US-08-447-570-28	Sequence 28, Appl
21	25.2	29.0	1723	2	US-08-643-763A-20	Sequence 20, Appl
22	25.2	29.0	1723	2	US-08-462-623-20	Sequence 20, Appl
23	25.2	29.0	1723	2	US-08-451-953A-20	Sequence 20, Appl
24	25.2	29.0	1723	3	US-08-459-346-5	Sequence 5, Appl
25	25.2	29.0	1723	3	US-08-445-468A-20	Sequence 20, Appl
26	25.2	29.0	1723	3	US-08-901-200A-7	Sequence 7, Appl
27	25.2	29.0	1723	3	US-08-449-700-28	Sequence 28, Appl
28	25.2	29.0	1723	4	US-08-449-699A-28	Sequence 28, Appl
29	25.2	29.0	1723	4	US-08-461-397A-20	Sequence 20, Appl
30	25.2	29.0	1723	4	US-08-912-088-20	Sequence 20, Appl
31	25.2	29.0	1723	5	US-08-278-730A-20	Sequence 20, Appl
32	25.2	29.0	1723	6	PCT-US92-01968-20	Sequence 20, Appl
33	25.2	29.0	1723	6	PCT-US92-05446-10	Sequence 10, Appl
34	25.2	29.0	1723	6	PCT-US93-07189-5	Sequence 5, Appl
35	25.2	29.0	1723	6	PCT-US93-07190-20	Sequence 20, Appl
36	25.2	29.0	1723	6	PCT-US93-07231-20	Sequence 20, Appl
37	25.2	29.0	1723	6	PCT-US93-08742-20	Sequence 20, Appl
38	25.2	29.0	1723	6	PCT-US93-08808-20	Sequence 20, Appl
39	25.2	29.0	1941	6	PCT-US93-08885-20	Sequence 20, Appl
40	25.2	29.0	1723	6	PCT-US91-07535-5	Sequence 5, Appl
41	25.2	29.0	6418	1	US-08-480-528A-11	Sequence 11, Appl
42	25.2	29.0	6418	1	US-08-479-666-11	Sequence 11, Appl
43	25.2	29.0	6418	5	PCT-US93-10520-11	Sequence 11, Appl
44	25.2	29.0	6418	5	PCT-US93-10520-11	Sequence 11, Appl
45	25	28.7	2467	4	US-08-701-240-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-167-919A-9
Sequence 9, Application US/0816791 9A
Patent No. 5674691
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,919A
FILING DATE: 18-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3132 base pairs
TYPE: nucleic acid

```
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 100..3048
US-08-167-919A-9

Qy    1 atggattgtcagctctccatcctcctccctccctccttccagctgctccggtcctcggacaagcttcggg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100 ATGGATTGTCAGCTCTCCATCCTCCTCCCTCCCTCCTTCCAGCTGCTCCGGTCCTCGGACAAGCTTCGGG

Qy   61 gaactgattccgcagccttccaatgaa 87
          |||||||||||||||||||||||||||
Db  160 GAACTGATTCCGCAGCCTTCCAATGAA 186


RESULT 2
US-08-715-106-9
; Sequence 9, Application US/08715106
; Patent No. 602030G
; GENERAL INFORMATION:
;   APPLICANT: Boyd, Andrew W.
;   APPLICANT: Simpson, Richard J.
;   APPLICANT: Wicks, Ian
;   APPLICANT: Ward, Larry D.
;   APPLICANT: Wilkinson, David
;   TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
;   TITLE OF INVENTION: AND USE THEREOF
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Scully, Scott, Murphy & Presser
;     STREET: 400 Garden City Plaza
;     CITY: Garden City
;     STATE: New York
;     COUNTRY: U.S.A.
;     ZIP: 11530
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/715,106
;     FILING DATE: 18-SEP-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/167,919
;     FILING DATE: 18-APR-1994
;     APPLICATION NUMBER: PK6841 (AU)
;     FILING DATE: 21-JUN-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PK9992 (AU)
;     FILING DATE: 12-DEC-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/AU92/00294
;     FILING DATE: 19-JUN-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: DiGiglio, Frank S.
;     REGISTRATION NUMBER: 31,346
;     REFERENCE/DOCKET NUMBER: 9159
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (516) 742-4343
;     TELEFAX: (516) 742-4366
;     TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 3132 base pairs
```

```
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 100..3048
US-08-715-106-9

Query Match             100.0%;  Score 87;  DB 1;  Length 3132;
Best Local Similarity   100.0%;  Pred. No. 4.6e-21;
Matches   87;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 atggattgtcagctctccatcctcctccctccctccttccagctgctccggtcctcggacaagcttcggg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100 ATGGATTGTCAGCTCTCCATCCTCCTCCCTCCCTCCTTCCAGCTGCTCCGGTCCTCGGACAAGCTTCGGG 159

Qy   61 gaactgattccgcagccttccaatgaa 87
          |||||||||||||||||||||||||||
Db  160 GAACTGATTCCGCAGCCTTCCAATGAA 186
```

```
RESULT 3
US-08-619-542B-34/c
; Sequence 34, Application US/0861954 2B
; Patent No. 5830662
; GENERAL INFORMATION:
;   APPLICANT: The Trustees of Columbia University in the City
;   APPLICANT: of New York
;   TITLE OF INVENTION: METHOD FOR CONSTRUCTION OF NORMALIZED
;   TITLE OF INVENTION: cDNA LIBRARIES
;   NUMBER OF SEQUENCES: 78
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Cooper & Dunham LLP
;     STREET: 1185 Avenue of the Americas
;     CITY: New York
;     STATE: New York
;     COUNTRY: USA
;     ZIP: 10036
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/619,542B
;     FILING DATE: June 21, 1996
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: White, John P.
;     REGISTRATION NUMBER: 28,678
;     REFERENCE/DOCKET NUMBER: 42840-A-PCT-US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) 278-0400
;     TELEFAX: (212) 391-0525
;     TELEX:
; INFORMATION FOR SEQ ID NO: 34:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 324 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 12..311
US-08-619-542B-34

Query Match             31.3%;  Score 27.2;  DB 3;  Length 324;
Best Local Similarity   67.9%;  Pred. No. 0.6;
Matches   38;  Conservative   0;  Mismatches   18;  Indels   0;  Gaps   0;
```

Qy   10 cagctctccatcctctcctcagctgctctgttctcgacagctgttctcgggggaact 65
         ||||||||||||||||||||||||||||||||||||||| || ||||||| ||||||
Db   260 CACCTCTTCATAATCCTTCTTCCAGAGCTGCCAGGTCCTCGCGCGGGGCCTCAGAGAACT 205

RESULT 4
5438126-1/c
;Patent No. 5438126
; TITLE OF INVENTION: HUMAN THYROID HORMONE RECEPTOR DNA
; APPLICANT: DEGROOT, LESLIE J.;NAKAI, AKIRA
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/830,766
;   FILING DATE: 03-FEB-1992
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 405,342
;   FILING DATE: 11-SEP-1989
;SEQ ID NO:1:
;   LENGTH: 1893

Query Match   29.9%; Score 26; DB 7; Length 1893;
Best Local Similarity   62.1%; Pred. No. 2.8;
Matches   41; Conservative   0; Mismatches   25; Indels   0; Gaps   0;

Qy   17 ccatctctcctcttctcagctgctctgttctcggggaactgattccgcagc 76
         | ||||| || |||||| |||||||||||||||||| || || || |||||
Db 1594 CCTTCCCTACCTTCTGAGAGGATGGGGTCCTGGGAACTGCAGGACACACAGCTCCCCAGC 1535

Qy   77 cttcca 82
         || |||
Db 1534 CTCCCA 1529

RESULT 5
US-08-619-542B-36/c
; Sequence 36, Application US/08619542B
; Patent No. 5830662
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City
; APPLICANT: of New York
; TITLE OF INVENTION: METHOD FOR CONSTRUCTION OF NORMALIZED
; TITLE OF INVENTION: cDNA LIBRARIES
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Cooper & Dunham LLP
;   STREET: 1185 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: USA
;   ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/619,542B
;   FILING DATE: June 21, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: White, John P.
;   REGISTRATION NUMBER: 28,678
;   REFERENCE/DOCKET NUMBER: 42840-A-PCT-US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 278-0400
;   TELEFAX: (212) 391-0525
;   TELEX:
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 318 base pairs

;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;US-08-619-542B-36

Query Match   29.4%; Score 25.6; DB 3; Length 318;
Best Local Similarity   66.1%; Pred. No. 2.1;
Matches   37; Conservative   0; Mismatches   19; Indels   0; Gaps   0;

Qy   10 cagctctccatcctctcctcagctgctctgttctcgacagctgttctcgggggaact 65
         ||||||||||||||||||||||||||||||||||||||| || ||||||| ||||||
Db   260 CACCTCTTCATAATCCTTCTTCTAGAGCTCCCAGGTCCTCGCGCGGGGCCTCAGAGAACT 205

RESULT 6
US-08-619-542B-37/c.
; Sequence 37, Application US/08619542B
; Patent No. 5830662
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City
; APPLICANT: of New York
; TITLE OF INVENTION: METHOD FOR CONSTRUCTION OF NORMALIZED
; TITLE OF INVENTION: cDNA LIBRARIES
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Cooper & Dunham LLP
;   STREET: 1185 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: USA
;   ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/619,542B
;   FILING DATE: June 21, 1996
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: White, John P.
;   REGISTRATION NUMBER: 28,678
;   REFERENCE/DOCKET NUMBER: 42840-A-PCT-US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 278-0400
;   TELEFAX: (212) 391-0525
;   TELEX:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..309
;US-08-619-542B-37

Query Match   29.4%; Score 25.6; DB 3; Length 323;
Best Local Similarity   57.5%; Pred. No. 2.1;
Matches   46; Conservative   0; Mismatches   34; Indels   0; Gaps   0;

Qy   2 tggattgtcagctcccatcctcctcctcagctgctctgttctctcgacagctgttcggggg 61
         | |||| | ||| ||| ||||| ||| ||||||||||||||||| ||||| ||| ||||
Db   275 TCGATGCCCACCTCCTCATAATCCTTCTTCTCCAGGCGCAGCCAATATCCTCACGGGGCCTCGGAG 216

Qy   62 aactgattccgcagcccttcc 81
         ||||| | || || ||| ||

Db   215 AACTCACCCTCCTCCATGCC 196

RESULT 7
US-08-901-200A-14
Sequence 14, Application US/08901200A
Patent No. 5854071
GENERAL INFORMATION:
APPLICANT: OPPERMANN, HERMANN
APPLICANT: OZKAYNAK, ENGIN
APPLICANT: KUBERASAMPATH, THANGAVEL
APPLICANT: RUEGER, DAVID C.
APPLICANT: PANG, ROY H.L.
APPLICANT: COHEN, CHARLES M.
TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
STREET: 45 SOUTH STREET
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 01748
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/901,200A
FILING DATE: 28-JUL-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: MEYERS, THOMAS C.
REGISTRATION NUMBER: 36,989
REFERENCE/DOCKET NUMBER: CRP-076DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 435-9001
TELEFAX: (508) 435-6951
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note= "A GAP OCCURS BETWEEN
OTHER INFORMATION: POSTION 884 IN SEQ ID NO: 11 AND POSITION 1 IN THIS SEQUENCE"
FEATURE:
NAME/KEY: exon
LOCATION: 204..393
OTHER INFORMATION: /note= "EXON TWO"
FEATURE:
NAME/KEY: exon
LOCATION: 466..930
OTHER INFORMATION: /note= "EXON THREE"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..950
OTHER INFORMATION: /note= "hOP-2 genomic sequence"

Qy   6 tgtcagctcctcatcctcctccaagctgctctgttctcgacagcttcggggaact 65
              |||| |||| ||||| |||| ||||||| ||| ||| || ||||| |||||| |||
Db   313 TTACAAGGTGCCCAGCATCCACCTGCTCAAACAGGACCCTCCACGTCACGTCAGCATGTTCCAGT 372

Query Match   29.0%; Score 25.2; DB 3; Length 950;
Best Local Similarity   57.7%; Pred. No. 4.2;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;
US-08-901-200A-14

Qy   66 gattccgcagcctccaa 83
              | |||| ||| || |||||
Db   373 GGTCCAGGAGCAGTCCAA 390

RESULT 8
US-07-800-364B-13
Sequence 13, Application US/07800364B
Patent No. 5688678
GENERAL INFORMATION:
APPLICANT: Hewick, Rodney M.
APPLICANT: Wang, Jack H.
APPLICANT: Wozney, John M.
APPLICANT: Celeste, Anthony J.
TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/800,364B
FILING DATE: 26-NOV-1991
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kaplnos, Ellen J.
REGISTRATION NUMBER: 32,245
REFERENCE/DOCKET NUMBER: GI 5182A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-876-1170
TELEFAX: 617-876-5851
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1003 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Human Heart
IMMEDIATE SOURCE:
LIBRARY: Human heart cDNA library stratagene catalog
CLONE: hh38
LIBRARY: #936208
POSITION IN GENOME:
UNITS: bp
FEATURE:
NAME/KEY: CDS
LOCATION: 8..850
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 427..843
FEATURE:
NAME/KEY: mRNA
LOCATION: 1..997

Query Match   29.0%; Score 25.2; DB 1; Length 1003;
Best Local Similarity   57.7%; Pred. No. 4.3;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;
US-07-800-364B-13

Qy   6 ttgtcagctctccatcctcctcctcagctgctctgttctcgacagcttcggggaact 65
              ||| ||| | ||| ||| ||| ||| | ||||| | ||||||||||| |
Db   88 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCCACGTCAGCATGTTCCAGGT 147

Qy   66 gattccgcagccttccaa 83
              ||| |||||| |||||
Db   148 GGTCCAGGAGGCAGTCCAA 165

RESULT 9
US-07-989-847-11
; Sequence 11, Application US/07989847
; Patent No. 5866364
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; APPLICANT: Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; TITLE OF INVENTION: Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
;   STREET: 87 CambridgePark Drive
;   CITY: Cambridge
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02140-2387
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Tape
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/989,847
;   FILING DATE:
;   CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kapinos, Ellen J.
;   REGISTRATION NUMBER: 32,245
;   REFERENCE/DOCKET NUMBER: GI-5192B
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-876-1170
;   TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1003 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
;   TISSUE TYPE: Human Heart
; IMMEDIATE SOURCE:
;   LIBRARY: Human heart cDNA library stratagene catalog
;   CLONE: #36208
; POSITION IN GENOME:
;   UNITS: bp
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 8..850
; FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 427..843
; FEATURE:
;   NAME/KEY: mRNA
;   LOCATION: 1..997

US-07-989-847-11

Query Match   29.0%; Score 25.2; DB 3; Length 1003;
Best Local Similarity   57.7%; Pred. No. 4.3;

Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy   6 ttgtcagctctccatcctcctcctcagctgctctgttctcgacagcttcggggaact 65
              ||| ||| | ||| ||| ||| ||| | ||||| | ||||||||||| |
Db   88 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCCACGTCAGCATGTTCCAGGT 147

Qy   66 gattccgcagccttccaa 83
              ||| |||||| |||||
Db   148 GGTCCAGGAGGCAGTCCAA 165

RESULT 10
US-07-841-646-28
; Sequence 28, Application US/07841646
; Patent No. 5266683
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: TESTA, HURWITZ & THIBEAULT
;   STREET: 53 STATE STREET
;   CITY: BOSTON
;   STATE: MASSACHUSETTS
;   COUNTRY: U.S.A.
;   ZIP: 02109
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/841,646
;   FILING DATE: 19920221
;   CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 810,560
;   FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 827,052
;   FILING DATE: 28-JAN-1992
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 660,162
;   FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 621,988
;   FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 621,849
;   FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 616,374
;   FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 600,024
;   FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 599,543
;   FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 579,865
;   FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 569,920
;   FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 483,913
;   FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 422,613
FILING DATE: 17-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 315,342
FILING DATE: 23-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 232,630
FILING DATE: 15-AUG-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 179,460
FILING DATE: 08-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-001CP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1723 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: HIPPOCAMPUS
FEATURE:
NAME/KEY: CDS
LOCATION: 490..1696
OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
OTHER INFORMATION: /product= "hOP2-PP"
OTHER INFORMATION: /note= "hOP2 (cDNA)"
US-07-841-646-28

Query Match   29.0%;   Score 25.2; DB 1; Length 1723;
Best Local Similarity 57.7%;   Pred. No. 5.1;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   6 ttgtcagctctccatcctcctccttcagctgctcgttctcgacagcttcgggggaact 65
           | |||| |||| || |||||||| || ||| | | | | || | |||
Db   933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCCTCCACGTCACCATGTTCCAGT 992

Qy   66 gattccgcagccttccaa 83
           ||| ||| |||| |||||
Db   993 GGTCCAGGAGCAGTCCAA 1010

RESULT 11
US-07-901-703-10
Sequence 10, Application US/07901703
Patent No. 5344654
GENERAL INFORMATION:
APPLICANT: RUEGER, DAVID C
APPLICANT: KUBERASAMPATH, THANGAVEL
APPLICANT: OPPERMANN, HERMANN
APPLICANT: OZAKAYNAK, ENGIN
TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED
OSTEOGENIC PROPERTIES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: EXCHANGE PLACE, 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/901,703
FILING DATE: 19920616
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: STK-057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1723 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: HIPPOCAMPUS
FEATURE:
NAME/KEY: CDS
LOCATION: 490..1696
OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
OTHER INFORMATION: /product= "hOP2-PP"
OTHER INFORMATION: /note= "hOP2 (cDNA)"
US-07-901-703-10

Query Match   29.0%;   Score 25.2; DB 1; Length 1723;
Best Local Similarity 57.7%;   Pred. No. 5.1;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   6 ttgtcagctctccatcctcctccttcagctgctcgttctcgacagcttcgggggaact 65
           | |||| |||| || |||||||| || ||| | | | | || | |||
Db   933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCCTCCACGTCACCATGTTCCAGT 992

Qy   66 gattccgcagccttccaa 83
           ||| ||| |||| |||||
Db   993 GGTCCAGGAGCAGTCCAA 1010

RESULT 12
US-08-147-023-28
Sequence 28, Application US/08147023
Patent No. 5468845
GENERAL INFORMATION:
APPLICANT: OPPERMANN, HERMANN
APPLICANT: OZKAYNAK, ENGIN
APPLICANT: KUBERASAMPATH, THANGAVEL
APPLICANT: RUEGER, DAVID C.
APPLICANT: PANG, ROY H.L.
TITLE OF INVENTION: OSTEOGENIC DEVICES
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/147,023
FILING DATE: 21-FEB-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 810,560

```
Qy    6 ttgtcagctctccatccctcctttccagctgtctgttctcgacaagcttcggggaact 65
          || || |  |||| |||| ||||   |||| ||| |||| || | ||  |||| ||||
Db  933 TTACAAGGTGCCCAGCATCCACCTGCTCACAGAGACCCTCCACGTCAGCATGTTCCAGGT 992

Qy   66 gattccgcagccttccaa 83
          | | ||||||  |||||
Db  993 GGTCCAGGAGGAGCAGTCCAA 1010
```

RESULT 13
US-08-206-864-3
Sequence 3, Application US/08206864
Patent No. 5610021
GENERAL INFORMATION:
  APPLICANT: RUEGER, DAVID C
  APPLICANT: JONES, WILLIAM K
  APPLICANT: TUCKER, RONALD F
  APPLICANT: OPPERMANN, HERMANN
  APPLICANT: OZKAYNAK, ENGIN
  APPLICANT: KUBERASAMPATH, THANGAVEL
  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
  TITLE OF INVENTION: RECOMBINANT OSTEOGENIC PROTEIN PRODUCTION
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
    ADDRESSEE: INC.
    STREET: 45 SOUTH STREET
    CITY: HOPKINTON
    STATE: MA
    COUNTRY: USA
    ZIP: 01748
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/206,864
    FILING DATE:
    CLASSIFICATION: 530
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/027,070
    FILING DATE: 04-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/841,646
    FILING DATE: 21-FEB-1992
  ATTORNEY/AGENT INFORMATION:
    NAME: PITCHER, EDMUND R
    REGISTRATION NUMBER: 27,829
    REFERENCE/DOCKET NUMBER: CRP-096
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617/7000
    TELEFAX: 617/248-7100
  INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1723 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    ORIGINAL SOURCE:
      ORGANISM: Homo sapiens
      TISSUE TYPE: HIPPOCAMPUS
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 490..1696
      OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
      OTHER INFORMATION: /product= "hOP2-PP"
      OTHER INFORMATION: /note= "hOP2 (cDNA)"
US-08-206-864-3

Query Match   29.0%; Score 25.2; DB 1; Length 1723;

FILING DATE: 20-DEC-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 827,052
  FILING DATE: 28-JAN-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 660,162
  FILING DATE: 22-FEB-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 621,988
  FILING DATE: 04-DEC-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 621,849
  FILING DATE: 04-DEC-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 616,374
  FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 600,024
  FILING DATE: 18-OCT-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 599,543
  FILING DATE: 18-OCT-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 579,865
  FILING DATE: 07-SEP-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 569,920
  FILING DATE: 20-AUG-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 483,913
  FILING DATE: 22-FEB-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 422,613
  FILING DATE: 17-OCT-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 315,342
  FILING DATE: 23-FEB-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 232,630
  FILING DATE: 15-AUG-1988
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 179,460
  FILING DATE: 08-APR-1988
ATTORNEY/AGENT INFORMATION:
  NAME: PITCHER, EDMUND R.
  REGISTRATION NUMBER: 27,829
  REFERENCE/DOCKET NUMBER: CRP-001CP6
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 617/248-7000
  TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 28:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1723 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  ORIGINAL SOURCE:
    ORGANISM: Homo sapiens
    TISSUE TYPE: HIPPOCAMPUS
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 490..1696
    OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
    OTHER INFORMATION: /product= "hOP2-PP"
    OTHER INFORMATION: /note= "hOP2 (cDNA)"
US-08-147-023-28

Query Match
Best Local Similarity 57.7%; Pred. No. 5.1;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

29.0%; Score 25.2; DB 1; Length 1723;
```

RESULT 14
US-08-278-729A-20

; Sequence 20, Application US/0827872 9A
; Patent No. 5650276
; GENERAL INFORMATION:
;  APPLICANT: SMART, JOHN
;  APPLICANT: OPPERMANN, HERMAN
;  APPLICANT: OZKAYNAK, ENGIN
;  APPLICANT: KUBERASAMPATH, THANGAVEL
;  APPLICANT: RUEGER, DAVID C.
;  APPLICANT: PANG, ROY H.L.
;  APPLICANT: COHEN, CHARLES M.
;  TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD
;  NUMBER OF SEQUENCES: 33
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
;  STREET: 45 SOUTH STREET
;  CITY: HOPKINTON
;  STATE: MA
;  COUNTRY: USA
;  ZIP: 01748
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/278,729A
;  FILING DATE: 20-JUL-1994
;  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;  NAME: PICHER Esq., EDMOND R.
;  REGISTRATION NUMBER: 27,829
;  REFERENCE/DOCKET NUMBER: CRP-058CPFW
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (508) 435-9001
;  TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 20:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1723 base pairs
;  TYPE: nucleic acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: cDNA
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: 490..1695
US-08-278-729A-20

Query Match            57.8%;  Score 25.2;  DB 1;  Length 1723;
Best Local Similarity  57.8%;  Pred. No. 5.1;
Matches   45;  Conservative  0;  Mismatches  33;  Indels  0;  Gaps  0;

Qy       6 ttgtcaagctctccatcctcctttcagctgctctgttccgacagcttcggggaact  65
           | ||| ||| |||||  || ||| |||| |||| | ||| ||||| ||||||| |
Db     933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCAACGTCAGCATGTTCCAGGT  992

Qy      66 gattccgcagcctccaa  83
           || |||||| ||||||
Db     993 GGTCCAGGAGGAGCAGTCCAA  1010

---

RESULT 15
US-08-480-528A-7

; Sequence 7, Application US/0848052 8A
; Patent No. 5652118
; GENERAL INFORMATION:
;  APPLICANT: OPPERMANN, HERMANN
;  APPLICANT: OZKAYNAK, ENGIN
;  APPLICANT: KUBERASAMPATH, THANGAVEL
;  APPLICANT: RUEGER, DAVID C.
;  APPLICANT: PANG, ROY H.L.
;  APPLICANT: COHEN, CHARLES M.
;  TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
;  NUMBER OF SEQUENCES: 13
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
;  STREET: 45 SOUTH STREET
;  CITY: HOPKINTON
;  STATE: MA
;  COUNTRY: USA
;  ZIP: 01748
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/480,528A
;  FILING DATE: 07-JUN-1995
;  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;  NAME: FENTON Esq., GILLIAN M.
;  REGISTRATION NUMBER: 36,508
;  REFERENCE/DOCKET NUMBER: CRP-076FW
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (617) 248-7560
;  TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 7:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1723 base pairs
;  TYPE: nucleic acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: cDNA
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: 490..1696
;  OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
;  OTHER INFORMATION: /product= "hOP2-PP"
;  OTHER INFORMATION: /note= "hOP2 cDNA"
US-08-480-528A-7

Query Match            29.0%;  Score 25.2;  DB 1;  Length 1723;
Best Local Similarity  57.8%;  Pred. No. 5.1;
Matches   45;  Conservative  0;  Mismatches  33;  Indels  0;  Gaps  0;

Qy       6 ttgtcaagctctccatcctcctttcagctgctctgttccgacagcttcggggaact  65
           | ||| ||| |||||  || ||| |||| |||| | ||| ||||| ||||||| |
Db     933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCAACGTCAGCATGTTCCAGGT  992

Qy      66 gattccgcagcctccaa  83
           || |||||| ||||||
Db     993 GGTCCAGGAGGAGCAGTCCAA  1010

Search completed: May 15, 2000, 12:05:03
Job time: 4422 sec

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          May 15, 2000, 12:08:19 ; Search time 111.14 Seconds
                 (without alignments)
                 195.849 Million cell updates/sec

Title:           US-09-104-340-6
Perfect score:   87
Sequence:        1 atggattgtcagctctccat..........ttccgcagccttccaatgaa 87

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters:        623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :       N_Geneseq_36:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match Length	Match DB Length	DB	ID	Description
	1	87	100.0	3132	1	Q34513	HEK coding sequenc
	2	27.2	31.3	1088	1	T76750	Ferritin cDNA. A p
	3	27.2	31.3	5562	1	V57542	Human calcium chan
c	4	27	31.0	110000	1	V21209_09	Continuation (10 o
c	5	26	29.9	1866	1	V84466	Human secreted pro
c	6	26	29.9	1893	1	Q95110	Human thyroid horm
c	7	25.6	29.4	360	1	X00663	Human secreted pro
c	8	25.4	29.2	1269	1	T94563	Human derived long
	9	25.4	29.2	1562	1	T94564	Human derived long
	10	25.2	29.0	1002	1	Q41296	Human BMP-8 gene.
	11	25.2	29.0	1723	1	Q28737	Human osteogenic p
	12	25.2	29.0	1723	1	Q38946	Human osteogenic p
	13	25.2	29.0	1723	1	Q38735	Human osteogenic p
	14	25.2	29.0	1723	1	Q38859	Morphogen hoP2 cod
	15	25.2	29.0	1723	1	Q53155	Sequence encoding
	16	25.2	29.0	1723	1	Q56200	hoP2 cDNA. Morphog
	17	25.2	29.0	1723	1	Q57917	hoP2 cDNA. Use mor
	18	25.2	29.0	1723	1	Q57313	Human OP-2. Morpho
	19	25.2	29.0	1723	1	Q45118	Human OP-2. Mainta
	20	25.2	29.0	1723	1	Q45164	Human OP-2. Use of
	21	25.2	29.0	1723	1	Q71426	Osteogenic protein
	22	25.2	29.0	1723	1	Q72426	hoP2-PP prepro for
	23	25.2	29.0	1723	1	T02600	hoP-2 cDNA. Antibo
	24	25.2	29.0	1723	1	T33443	Human osteogenic P
	25	25.2	29.0	1723	1	T73208	Human osteogenic P
	26	25.2	29.0	1723	1	T97881	Human hippocampal
	27	25.2	29.0	1723	1	V10347	cDNA encoding huma
	28	25.2	29.0	1723	1	V15218	Human OP-2 cDNA. T
	29	25.2	29.0	1723	1	V19535	Human osteogenic p
	30	25.2	29.0	1723	1	V32585	Human osteogenic p
	31	25.2	29.0	1723	1	X00243	Human osteogenic p
	32	25.2	29.0	1941	1	O24522	cDNA encoding prep
c	33	25.2	29.0	2358	1	V45134	Mouse Fas-binding
	35	25.2	29.0	6418	1	O65395	Osteogenic protein
	36	25	28.7	2467	1	T61862	Human glial tumour
	37	25	28.7	6327	1	T32301	Dermatomyositis sp
	38	24.6	28.3	2883	1	T58898	C-Delta-1 gene (al
c	39	24.6	28.0	244	1	X39537	Human secreted pro
c	40	24.4	28.0	4931	1	V29584	Homo sapiens sulph
	41	24.4	28.0	2658	1	T60039	Pumpkin ent-kauren
	42	24.	27.6	3225	1	T15224	Tumor necrosis fac
c	43	24	27.6	3225	1	T94637	TNF-R1-DD ligand p
c	44	24	27.6	6373	1	V74438	Staphylococcus aur
c	45	23.8	27.4	875	1	X39853	Gastric cancer ass

ALIGNMENTS

RESULT 1

```
ID   Q34513  standard; DNA; 3132 BP.
AC   Q34513;
DT   24-MAY-1993  (first entry)
DE   HEK coding sequence.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FH
FT   5'utr           1..99
FT                   /*tag= a
FT   cds             100..3051
FT                   /*tag= b
FT   3'utr           3052..3132
FT                   /*tag= c
FT   signal_peptide  100..159
FT                   /*tag= g
FT                   /note= "Purified HEK protein #2"
FT   misc_feature    1723..1795
FT                   /*tag= d
FT   mat_peptide     160..216
FT                   /*tag= f
FT                   /note= "Potential transmembrane region"
FT   mat_peptide     2617..2679
FT                   /*tag= e
FT                   /note= "Purified HEK protein #1"
PN   WO9300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU0294.
PR   21-JUN-1991; AU-0006841.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD; Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   P-PSDB; R31466.
DR   WPI; 93-036373/04.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence encodes human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the
CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence 3132 BP; 888 A; 711 C; 759 G; 774 T;

Query Match             100.0%;  Score 87;  DB 1;  Length 3132;
Best Local Similarity   100.0%;  Pred. No. 3.9e-19;
Matches   87;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy    1 atggatgtgcaagctctccatcctcctccttctcaagctgtctgttctcgacagcttcggg  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  100 ATGGATGTGCAGCTGTCAGCTCTCCATCCTCCTCCTTCTCAAGCTGTCTGTTCTCGACAGCTTCGGG 159

Qy   61 gaactgattccgcagcctccaatgaa  87
        ||||||||||| ||||||| |||||||
Db  160 GAACTGATTCCCGCAGCCTTCCAATGAA 186
```

RESULT 2
T76750
```
ID   T76750; standard; cDNA to mRNA; 1088 BP.
AC   T76750;
DT   08-OCT-1997 (first entry)
DE   Ferritin cDNA.
KW   Plant; ferritin; acidic soil; diagnosis; agriculture; ds.
OS   Acacia mangium.
FH   Key             Location/Qualifiers
FT   cds             71..877
FT                   /tag= a
FT                   /product= Ferritin
FT   transit_peptide 71..247
FT                   /tag= b
FT   mat_peptide     248..874
FT                   /tag= c
FT   polya_signal    1055..1059
FT                   /tag= d
FT   misc_feature    1079..1088
FT                   /tag= e
FT                   /note= "PolyA_site"
PN   J09140384-A.
PD   03-JUN-1997.
PF   20-NOV-1995; 323527.
PR   20-NOV-1995; JP-323527.
PA   (NETT-) NETTAI RIN SAISEI GIJUTSU KENKYU KUMIAI.
DR   WPI: 97-344897/32.
DR   P-PSDB; W21736.
PT   A plant ferritin gene induced in acidic soil - derived from Acacia
     plants, improves resistance to acidic growth conditions
PS   Claim 1; Page 6-7; 9pp; Japanese.
CC   This sequence represents a plant ferritin gene. The plant ferritin
CC   gene, whose production is induced in a plant caused by acidic soil, is involved in the
CC   occurrence of disorders in a plant caused by acidic soil as well as
CC   in improvements in the resistance of a plant to acidic conditions.
CC   This gene can be used for breeding of plants highly resistant to acidic
CC   soil. The cDNA or its fragment can also be used for diagnosis of a
CC   plant planted in acidic soil by determining the expression of ferritin
CC   in it. Ferritin can be expressed excessively by introducing the cDNA
CC   of the invention into a plant in the sense direction to reinforce the
CC   resistance of the plant to acidic soil. Therefore acidic soil which could
CC   not be utilised for agricultural purposes can be utilised by the
CC   resulting plants.
SQ   Sequence 1088 BP;   273 A;   221 C;   254 G;   340 T;
```

```
Query Match   31.38;  Score 27.2;  DB 1;  Length 1088;
Best Local Similarity   64.18%;                 Pred. No. 3.9;
Matches   41;  Conservative   0;  Mismatches   23;  Indels   0;  Gaps   0;

Qy   16 tccatcctcctccttctcaagctgctcttgttctcgacaagcttcggggaactgattccgcag  75
        || ||||| ||||| |||| ||| ||||||||||||||| |||||| ||
Db    1 TTCATCTTCTTCTTCTTCTTCTTTCAATTCAATTCTTCTTCCGGAGTTCGGTTCCTTGAG  60

Qy   76 cctg  79
        | |||
Db   61 CTTT  64
```

```
DT   20-NOV-1998 (first entry)
DE   Human calcium channel subunit alpha-1 partial sequence encoding cDNA.
KW   Calcium channel subunit alpha-1; recombinant; alpha 2 delta;
KW   beta calcium channel; agonist; antagonist; alpha-1I calcium channel;
KW   histological assay; tissue distribution; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
CDS              1..5562
FT                   /tag= a
FT                   /product= "Human calcium channel subunit alpha-1
FT                   partial sequence"
PN   WO9838301-A1.
PD   03-SEP-1998.
PF   27-FEB-1998; CA0173.
PR   25-FEB-1998; US-030482.
PR   28-FEB-1997; US-030204.
PA   (NEUR-) NEUROMED TECHNOLOGIES INC.
PI   Baillie DL; Snutch TP;
DR   WPI: 98-481203/41.
DR   P-PSDB; W79161.
PT   Newly identified calcium channel sub:units alpha 1I and alpha 1H -
     for development of cell lines which express the novel calcium
     channels, useful for screening channel (ant)agonists
PS   Claim 2; Pages 26-33; 45pp; English.
CC   This represents a cDNA from BAC bk206c7 encoding a partial sequence of a
CC   human calcium channel subunit alpha-1. The invention provides isolated
CC   DNA fragments coding for novel calcium channel subunits alpha-1I and
CC   alpha-1H. An eukaryotic cell transiently or stably transformed with an
CC   expression vector containing the calcium subunits encoding DNA fragments
CC   can be used for expressing the calcium channel. The cells are optionally
CC   further transformed to express alpha 2 delta or beta calcium channel or
CC   both the proteins. The transformed cells are useful for identifying
CC   compounds capable of acting as agonists or antagonists for the alpha-1I
CC   calcium channel. The nucleic acid sequences can be used in histological
CC   assay to determine the tissue distribution of the novel calcium channel
CC   subunits.
SQ   Sequence 5562 BP;   1060 A;   1757 C;   1654 G;   1091 T;
```

```
Query Match   31.38;  Score 27.2;  DB 1;  Length 5562;
Best Local Similarity   72.9%;                  Pred. No. 5.6;
Matches   35;  Conservative   0;  Mismatches   13;  Indels   0;  Gaps   0;

Qy   14 tctccatcctcctccttctcaagctgctctgttctcgacaagcttcggg  61
        ||||| ||| || ||| ||||||||||||| ||||| ||||||| |||
Db 5258 TCTCCTTCTGCTCATCGTCAGCCTTCTTGTGCTCAACATGTTGTGG 5305
```

RESULT 4
V21209_09/c
Continuation (10 of 17) of V21209 from base 900001 (Methanococcus Jannaschii circular
WP Sequence split into 17 fragments LOCUS V21209 Accession V21209

WP Fragment Name	Begin	End
V21209_00	1	110000
V21209_01	110001	210000
V21209_02	200001	310000
V21209_03	300001	410000
V21209_04	400001	510000
V21209_05	500001	610000
V21209_06	600001	710000
V21209_07	700001	810000
V21209_08	800001	910000
V21209_09	900001	1010000
V21209_10	1000001	1110000
V21209_11	1100001	1210000
V21209_12	1200001	1310000
V21209_13	1300001	1410000
V21209_14	1400001	1510000
V21209_15	1500001	1610000
V21209_16	1600001	1664976

```
Query Match   31.0%;  Score 27;  DB 1;  Length 110000;
```

RESULT 3
V57542
```
ID   V57542; standard; cDNA; 5562 BP.
AC   V57542;
```

Query Match

```
Best Local Similarity  60.0%;  Pred. No. 12;
Matches   45;  Conservative   0;  Mismatches   30;  Indels   0;  Gaps   0;

Qy        5 attgtcagctcctccatcctcctcctcagctgtcctctcgacagcttcgggaac  64
            ||||| |||| |   ||||| |  ||||||||  || ||||  | |||||| |||
Db    33054 ATTGGATTCTCTACATTCAAAAACCTTCTCTCCTTTTTTCTTTAGCTRAAATCTTCCAATAAC  32995

Qy       65 tgattccgcagcctt  79
            |||||  ||||||||
Db    32994 TCATTCAATAGCTTT  32980


RESULT   5

ID   V84466 standard; DNA; 1866 BP.
AC   V84466;
DT   01-MAR-1999 (first entry)
DE   Human secreted protein gene 56 clone HE2OF09.
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS   Homo sapiens.
PN   WO9854963-A2.
PD   10-DEC-1998.
PF   04-JUN-1998; U11422.
PR   18-DEC-1997; US-070923.
PR   01-MAR-1999; US-048877.
PR   06-JUN-1997; US-048881.
PR   06-JUN-1997; US-048884.
PR   06-JUN-1997; US-048893.
PR   06-JUN-1997; US-048896.
PR   06-JUN-1997; US-048899.
PR   06-JUN-1997; US-048915.
PR   06-JUN-1997; US-048949.
PR   06-JUN-1997; US-048964.
PR   06-JUN-1997; US-048972.
PR   06-JUN-1997; US-049020.
PR   06-JUN-1997; US-049375.
PR   05-SEP-1997; US-057628.
PR   05-SEP-1997; US-057635.
PR   05-SEP-1997; US-057644.
PR   05-SEP-1997; US-057647.
PR   05-SEP-1997; US-057650.
PR   05-SEP-1997; US-057661.
PR   05-SEP-1997; US-057667.
PR   05-SEP-1997; US-057761.
PR   05-SEP-1997; US-057764.
PR   05-SEP-1997; US-057770.
PR   05-SEP-1997; US-057775.
PR   05-SEP-1997; US-057778.
PR   06-JUN-1997; US-048875.
PR   06-JUN-1997; US-048878.
PR   06-JUN-1997; US-048882.
PR   06-JUN-1997; US-048885.
PR   06-JUN-1997; US-048894.
PR   06-JUN-1997; US-048897.
PR   06-JUN-1997; US-048900.
PR   06-JUN-1997; US-048916.
PR   06-JUN-1997; US-048962.
PR   06-JUN-1997; US-048970.
PR   06-JUN-1997; US-048974.
PR   06-JUN-1997; US-049373.
PR   05-SEP-1997; US-057584.
PR   05-SEP-1997; US-057629.
PR   05-SEP-1997; US-057642.
PR   05-SEP-1997; US-057645.
PR   05-SEP-1997; US-057648.
PR   05-SEP-1997; US-057651.
PR   05-SEP-1997; US-057662.
PR   05-SEP-1997; US-057668.
PR   05-SEP-1997; US-057762.
PR   05-SEP-1997; US-057765.
PR   05-SEP-1997; US-057771.
PR   06-JUN-1997; US-049374.
PR   05-SEP-1997; US-057627.
PR   05-SEP-1997; US-057634.
PR   05-SEP-1997; US-057643.
PR   05-SEP-1997; US-057646.
PR   05-SEP-1997; US-057649.
PR   05-SEP-1997; US-057654.
PR   05-SEP-1997; US-057666.
PR   05-SEP-1997; US-057760.
PR   05-SEP-1997; US-057763.
PR   05-SEP-1997; US-057769.
PR   05-SEP-1997; US-057774.
PR   05-SEP-1997; US-057777.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
PI   Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,
PI   Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW,
PI   Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI   Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
DR   WPI; 99-059865/05.
DR   P-PSDB; W88589, W88807, W88808, W88809, W88810, W88811,
DR   P-PSDB; W88534 to W88756.
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
PS   Claim 4; Page 323-324; 772pp; English.
CC   The invention relates to nucleic acid sequences (V84411 to V84633)
CC   encoding human secreted proteins (W88534 to W88756). The secreted protein
CC   gene sequences are deposited with the ATCC under deposit numbers ATCC
CC   97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC   209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC   cells comprising recombinant vectors containing the nucleic acid
CC   sequences are used for the recombinant production of the secreted
CC   proteins. The polynucleotide and amino acid sequences are useful for are
CC   useful for preventing, treating or ameliorating medical conditions e.g.
CC   by protein or gene therapy. Pathological conditions can be also diagnosed
CC   by determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the polynucleotides, based on
CC   determining the amount of the new polypeptides in a sample or by
CC   which tissues they are most highly expressed in, and include developing
CC   products for the diagnosis or treatment of cancer, neurodegenerative
CC   disorders, developmental abnormalities and foetal deficiencies, blood
CC   disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC   diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC   ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC   restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC   such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC   or thymus, digestive/endocrine disorders, infections and AIDS. The
CC   polypeptides are also useful for identifying their binding partners. The
CC   present sequence represents a gene encoding a human secreted protein
CC   (see descriptor line for gene number and clone identification).
SQ   Sequence 1866 BP;  484 A;  458 C;  397 G;  524 T;

Query Match   29.9%;  Score 26;  DB 1;  Length 1866;
Best Local Similarity   59.5%;  Pred. No. 11;
Matches   44;  Conservative   0;  Mismatches   30;  Indels   0;  Gaps   0;
```

```
Qy      1 atggattgtcagctccatcctccctccttctcagctgctctcgtttcttcgacagcttcggg 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105 AAGGCAtTCtAACtCTTCCAtTCGTCCTACATCCAAGAtGCTTTGTGTCTtCACAGGCCGTGAG 46

Qy     61 gaactgattccgca 74
          ||||||||||||
Db     45 GAACAGCtTTGCA 32
```

RESULT   6

ID   Q95110/c
AC   Q95110;
DT   06-MAR-1996 (first entry)
DE   Human thyroid hormone receptor alpha-1 cDNA.
PN   US543126-A.
PD   01-AUG-1995.
PF   11-SEP-1989; 405342.
PR   11-SEP-1989; US-405342.
PR   03-FEB-1992; US-830766.
PA   (ARCH-) ARCH DEV CORP.
PI   Degroot LJ, Nakai A;
DR   WPI; 95-274923/36.
DR   P-PSDB; R78318.

CC   New human thyroid hormone receptor nucleic acid - used to develop
CC   prods. for use in analysis, assays and therapeutic agents
PS   Claim 1; Fig 1; 7pp; English.
CC   Q95110 encodes R78318 human thyroid hormone receptor alpha-1 (hTR-
CC   alpha-1).The nucleic acid can be used to develop prods. for use in
CC   analysis assays and therapeutic agents. hTR-alpha-1 can be used in
CC   tests for thyroid function, or for producing antibodies for use in
CC   affinity purifcn., detection and quantification.
SQ   Sequence 1893 BP; 461 A; 531 C; 551 G; 350 T;

```
Query Match             29.9%;  Score 26;  DB 1;  Length 1893;
Best Local Similarity   62.1%;  Pred. No. 11;
Matches    41;  Conservative    0;  Mismatches    25;  Indels    0;  Gaps    0;

Qy     17 ccatcctccctccttctcagctgctctcgtttcttcgacagcttcggggaactgattccgcagc 76
          |||| ||||||| ||||||| |||| ||||||||||||||||||||||||||| ||||
Db   1594 CCTtCCCTACCTTTCTGAGAGGAtNGGGTCCTGGGAACTGCAGGACACAGCtTCCCCCAGC 1535
```

RESULT   7

ID   X00663
AC   X00663;
DT   25-MAR-1999 (first entry)
DE   Human secreted protein gene 53 clone HBMCL41.
KW   Human; secreted protein; fusion protein; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; restenosis;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS   Homo sapiens.
PN   WO9842738-A1.
PD   01-OCT-1998.
```
FH    Key             Location/Qualifiers
FT    cds             73..1305
FT                    /*tag= a
```

PF   19-MAR-1998;   U05311.
PR   30-MAY-1997;   US-050937.
PR   21-MAR-1997;   US-0411276.
PR   21-MAR-1997;   US-0411277.
PR   21-MAR-1997;   US-041281.
PR   21-MAR-1997;   US-042344.
PR   21-MAR-1997;   US-048069.
PR   30-MAY-1997;   US-048094.
PR   30-MAY-1997;   US-048095.
PR   30-MAY-1997;   US-048096.
PR   30-MAY-1997;   US-048099.
PR   30-MAY-1997;   US-0481131.
PR   30-MAY-1997;   US-048154.
PR   30-MAY-1997;   US-048160.
PR   30-MAY-1997;   US-048187.
PR   30-MAY-1997;   US-048186.
PR   30-MAY-1997;   US-048188.
PR   30-MAY-1997;   US-048350.
PR   30-MAY-1997;   US-048351.
PR   30-MAY-1997;   US-048352.
PR   30-MAY-1997;   US-048355.
PR   05-AUG-1997;   US-054804.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Brewer LA, Duan R, Ferrie AM, Florence KA,
PI   Greene JM, Hu JS, Ebner R, Moore PA, Ni J, Olsen HS,
PI   Rosen CA, Ruben SM, Shi Y, Young P;
DR   WPI; 99-070006/06.
DR   P-PSDB; W67859.

CC   New isolated human genes and the secreted polypeptides they encode -
CC   useful for diagnosis and treatment of e.g. cancers, neurological
CC   disorders, immune diseases, inflammation or blood disorders
PS   Claim 1; Page 215; 385pp; English.
CC   This sequence represents a nucleic acid molecule which encodes a secreted
CC   human protein. The gene number and the clone it is derived from, are
CC   detailed in the descriptor line. The gene can be used to generate fusion
CC   proteins by linking to the gene to a human immunoglobulin Fc portion
CC   (e.g. X00602) for increasing the stability of the fused protein as
CC   compared to the human protein only.
PT   The invention relates to 87 novel genes and their fragments (nucleic acid
CC   sequences: X00611-X00724; amino acid sequences W67807-W68004) which
CC   are useful for preventing, treating or ameliorating medical conditions
CC   e.g. by protein or gene therapy. Also, pathological conditions can be
CC   diagnosed by determining the presence of the new polypeptides in a sample
CC   or by determining the amount of the new polypeptides. In a sample
CC   specific uses are described for each of the 87 polynucleotides, based on
CC   which tissues they are most highly expressed in (see X00611 for described
CC   uses).
SQ   Sequence 360 BP; 102 A; 89 C; 76 G; 91 T;

```
Query Match             29.4%;  Score 25.6;  DB 1;  Length 360;
Best Local Similarity   63.8%;  Pred. No. 9.9;
Matches    37;  Conservative    1;  Mismatches    20;  Indels    0;  Gaps    0;

Qy     30 tctcagctgctctcgtttcttcgacagcttcggggaactgattccgcagccttcaaatgaa 87
          ||| || |||| ||||||||||||||||||||||||||||||||||||||||| |||
Db     44 TGTAAGGCCCGCAGCTCCCGCGACGCTCCCGCCGGACTCSTGCGCCGCCTCCTTACCATGAA 101
```

RESULT   8

ID   T94563
AC   T94563;
DT   22-APR-1998 (first entry)
DE   Human derived long chain DNA novel gene.
KW   Human; long chain DNA; antibody; nervous disorder; detection;
KW   inhibitor; ss.
OS   Homo sapiens.
```
FH    Key             Location/Qualifiers
FT    CDS             1..1269
FT                    /*tag= a
FT                    /product= "long chain DNA protein"
```

```
FT                    /note= "no stop codon given"
PN   WO9737018-A1.
PD   09-OCT-1997.
PF   19-MAR-1997; J00919.
PR   03-APR-1996; JP-081304.
PA   (ASAH ) ASAHI KASEI KOGYO KK.
PI   Kato C; Takada Y;
DR   WPI; 97-503103/46.
DR   P-PSDB; W32797.
PT   Human long chain DNA and antibodies which bind to the peptides(s) -
PT   is potentially useful in treating nervous disorders
PS   Claim 3; Page 91-92; 109pp; Japanese.
CC   The present sequence represents human derived long chain DNA comprising
CC   at least a peptide-encoding region. Peptides encoded by the human
CC   derived long chain DNA, may be used for detecting substances which bind
CC   to the peptides, and for detecting substances which suppress or inhibit
CC   binding to these peptides. The peptides and their inhibitors are
CC   potentially useful in treating nervous disorders.
SQ   Sequence 1269 BP;    263 A;    413 C;    285 G;    308 T;

Qy          10  cagctctccatcctcctctcagctgctcctgttctcgacaagcttcggggaactgatt  69
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db         736  CTGCTCTACAACCTGCCCCCTCCTCACCATCTCTGTGCCCTACGCTCTGTGTGGCCAAGAAA  795

Qy          70  ccgcagccttccaat  84
                ||| ||||| | |||
Db         796  CTGTGGCTGTGTAAT  810


RESULT  9
ID   T94564 standard; DNA; 1562 BP.
AC   T94564;
DT   22-APR-1998 (first entry)
DE   Human derived long chain DNA novel gene.
KW   Human; long chain DNA; antibody; nervous disorder; detection;
KW   inhibitor; ss.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   CDS              168..1439
FT                    /tag=  a
FT                    /product= "long chain DNA protein"
PN   WO9737018-A1.
PD   09-OCT-1997.
PF   19-MAR-1997; J00919.
PR   03-APR-1996; JP-081304.
PA   (ASAH ) ASAHI KASEI KOGYO KK.
PI   Kato C; Takada Y;
DR   WPI; 97-503103/46.
DR   P-PSDB; W32797.
PT   Human long chain DNA and antibodies which bind to the peptides(s) -
PT   is potentially useful in treating nervous disorders
PS   Claim 4; Page 92-93; 109pp; Japanese.
CC   The present sequence represents human derived long chain DNA comprising
CC   at least a peptide-encoding region. Peptides encoded by the human
CC   derived long chain DNA, may be used for detecting substances which bind
CC   to the peptides, and for detecting substances which suppress or inhibit
CC   binding to these peptides. The peptides and their inhibitors are
CC   potentially useful in treating nervous disorders.
SQ   Sequence 1562 BP;    318 A;    506 C;    389 G;    349 T;

Query Match             29.2%;  Score 25.4;  DB 1;  Length 1562;
Best Local Similarity   58.7%;  Pred. No. 16;
Matches    44;  Conservative    0;  Mismatches    31;  Indels    0;  Gaps    0;

Qy          10  cagctctccatcctcctctcagctgctcctgttctcgacaagcttcggggaactgatt  69
                ||||||| || ||| |||  || | || ||| || ||  | |||||  | |||| |||
Db        1562  ...
```

Right column:

```
Db         903  CTGCTCTACACATCCTGCCCCCTCCTCCCATCATCTCTGTGCCCTAGGCTCTGGTGGCCAAGAAA  962

Qy          70  ccgcagccttccaat  84
                ||| ||||| | |||
Db         963  CTGTGGCTGTGTAAT  977


RESULT  10
ID   Q41296 standard; DNA; 1002 BP.
AC   Q41296;
DT   13-SEP-1993 (first entry)
DE   Human BMP-8 gene.
KW   Bone morphogenetic protein; bone defect treatment; healing; wound;
KW   injury; tissue repair; osteoporosis; burns; incisions; ulcers;
KW   neuronal survival increase; fracture reduction; cartilage growth;
KW   induction; ss.
OS   Homo sapiens.
FH   Key              Location/Qualifiers
FT   cds              8..850
FT                    /tag=  a
FT   mat_peptide      427..843
FT                    /tag=  b
FT   mrna             1..997
FT                    /tag=  c
PN   WO9309229-A.
PD   13-MAY-1993.
PF   02-NOV-1992; U09430.
PR   04-NOV-1991; US-787496.
PR   07-APR-1992; US-864692.
PA   (GEMY ) GENETICS INST INC.
PI   Israel D, Wolfman NM;
DR   WPI; 93-167695/20.
DR   P-PSDB; R36737.
PT   Recombinant hetero-dimeric BMP proteins - are useful in treating
PT   bone defects, healing bone injury and in wound healing
PS   Disclosure; Fig 6; 16pp; English.
CC   The sequence is that encoding the human bone morphogenetic protein
CC   BMP-8. It may be used in the prodn. of a recombinant heterodimeric
CC   protein having bone stimulating activity. This heterodimer is
CC   encoded by a sequence encoding BMP-2 or BMP-4 or a fragment and a
CC   sequence encoding a second protein or fragment, BMP-8. It may be
CC   used in compsns. for wound healing, tissue repair, and in similar
CC   compsns. which have been indicated for the use of individual BMPs.
CC   Increased potency of the heterodimer over individual BMPs may permit
CC   lower dosages to be administered. A heterodimeric protein which
CC   induces cartilage and/or bone growth in circumstances where bone
CC   is not normally formed, has applications in the healing of bone
CC   fractures and cartilage defects in humans and other animals. The
CC   heterodimer may have prophylactic use in closed as well as open
CC   fracture reduction and also in the improved fixation of artificial
CC   joints. De novo bone formation induced by an osteogenic agent
CC   contributes to the repair of congenital, trauma induced or oncologic
CC   resection induced craniofacial defects, and also is useful in cosmetic
CC   plastic surgery. It may be used in the treatment of periodontal
CC   disease and in other tooth repair processes. It may also be useful
CC   in the treatment of osteoporosis, wound healing (e.g. burns,
CC   incisions and ulcers) and related tissue repair, and may increase
CC   neuronal survival and be useful in the transplantation and treatment
CC   of conditions exhibiting a decrease in neuronal survival. It may be
CC   combined with other agents beneficial to the bone and/or cartilage
CC   defect, wound or tissue in question, e.g. EGF, PDGF, TGF-alpha,
CC   TGF-beta and insulin-like growth factor
SQ   Sequence 1002 BP;    196 A;    330 C;    285 G;    191 T;

Query Match             29.0%;  Score 25.2;  DB 1;  Length 1002;
Best Local Similarity   57.7%;  Pred. No. 17;
Matches    45;  Conservative    0;  Mismatches    33;  Indels    0;  Gaps    0;

Qy           6  ttgtcagctctccatcctcctctcagctgctcctgttctcgacaagcttcggggaact  65
                 || ||| ||| || |||  | ||  | || ||| ||  | |||||  | |||| ||
Db          88  TTACAAGGTGCCCAGCATCCACCTGCTCACAACAGGACCCTCCACGTCAGCATGTTCCAGGT  147
```

```
Qy    66 gattccgcagccttccaa 83
         |||||||||| |||||
Db   148 GGTCCAGGAGCAGTCCAA 165
```

RESULT 11
ID   Q28737
     Q28737 standard; cDNA; 1723 BP.
AC   Q28737;
DT   26-FEB-1993 (first entry)
DE   Human osteogenic protein hoP2 coding sequence.
KW   Morphogen; morphogenic protein.
OS   Homo sapiens.
FH
FT   Key   Location/Qualifiers
FT   cds   490..1698
FT   /*tag= a
FT   /product= hoP2

PN   WO921523-A.
PD   17-SEP-1992;
PF   11-MAR-1992; US-667274.
PR   11-MAR-1991; US-667274.
PA   (CREA-) CREATIVE BIOMOLECULES INC.
PI   Cohen CM, Kuberasampath T, Oppermann H, Pang RHL, Rueger DC;
DR   WPI; 92-331475/40.
DR   P-PSDB; R27291.
PT   Compsns. for increasing progenitor cell population - contain a
PT   morphogen to induce proliferation, useful for inhibiting
PT   neoplastic growth, inducing tissue repair and in diagnosis of
PT   tissue dysfunction
PS   Disclosure; Page 102-104; 132pp; English.
CC   Mature hoP2 is one of the preferred known morphogens which can be
CC   used in the manufacture of pharmaceuticals for inducing non-
CC   chondrogenic mammalian tissue growth, progenitor cell proliferation
CC   and heaptic tissue growth and for maintaining the phenotypic
CC   expression of differentiated cells in a mammal. Morphogenic
CC   compositions of the invention can also be used to treat blood
CC   disorders and impaired or lost immune function. Morphogens sharing
CC   at least 70% homology with hoP2 are included.
SQ   Sequence 1723 BP; 265 A; 624 C; 576 G; 258 T;

Query Match   29.0%; Score 25.2; DB 1; Length 1723;
Best Local Similarity   57.7%;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy    6 ttgtcagctccatcctcctccttctcagctgctctcgttctcgacagcttcggggaact 65
        ||| || ||| | ||| || ||| |||||| || | ||| ||| |||
Db  933 TTACAAGGTGCCCAGCATCCACCTGCTCAAACAGGACCCTCACGTCAGCATGTTCCAGT 992

Qy   66 gattccgcagccttccaa 83
        |||||||||| |||||
Db  993 GGTCCAGGAGCAGTCCAA 1010
```

RESULT 12
ID   Q38946
     Q38946 standard; cDNA; 1723 BP.
AC   Q38946;
DT   21-JUL-1993 (first entry)
DE   Human osteogenic protein 2 (hoP-2) gene.
KW   Bone; loss; increase; fracture; post-menopausal; senile;
KW   osteoporosis; hyperparathyroidism; skeletal microstructure defects;
KW   chronic renal failure; kidney disease; osteomalacia, vitamin D;
KW   deficiency-induced osteopenia; osteoporosis; Paget's disease;
KW   bone mass; imbalance; resorption; formation; dialysis; calcium;
KW   phosphate; metabolism; ss.
OS   Homo sapiens.
FH
FT   Key   Location/Qualifiers
FT   cds   490..1696
FT   /*tag= a
FT   /note= "hoP-2"

PN   WO9305751-A.

PD   01-APR-1993.
PF   28-AUG-1992; U07432.
PF   30-AUG-1991; US-752764.
PR   30-AUG-1991; US-752857.
PR   30-AUG-1991; US-752861.
PR   31-JUL-1992; US-923780.
PA   (CREA-) CREATIVE BIOMOLECULES INC.
PI   Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E,
PI   Pang RHL, Rueger DC, Smart JE;
DR   WPI; 93-117208/14.
PT   Use of morphogen or in-vivo morphogenic-stimulating agent - to
PT   prevent bone loss or increase, used for treating bone fractures,
PT   post-menopausal or senile osteoporosis, hyperparathyroidism etc.
PS   Disclosure; Page 115-117; 162pp; English.
CC   The sequence is that encoding human osteogenic protein 2 (hoP-2) a
CC   a morphogenically active protein which may be used as part of a
CC   method for treating a bone fracture or a disease which causes or
CC   results in bone fractures or other defects in skeletal
CC   microstructure. Such diseases include chronic renal failure and
CC   other kidney diseases, osteomalacia, vitamin D deficiency-induced
CC   osteopenia or osteoporosis, postmenopausal or senile osteoporosis,
CC   hyperparathyroidism and Paget's disease. The methods can be used for
CC   protecting individuals at risk for loss of bone mass such as
CC   postmenopausal females, aged individuals and individuals undergoing
CC   dialysis. The loss of bone mass may result from an imbalance in bone
CC   resorption or bone formation, an imbalance of calcium or phosphate
CC   metabolism, a vitamin D imbalance or be nutritionally or hormonally
CC   induced.
SQ   Sequence 1723 BP; 266 A; 624 C; 576 G; 257 T;

Query Match   29.0%; Score 25.2; DB 1; Length 1723;
Best Local Similarity   57.7%;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy    6 ttgtcagctccatcctcctccttctcagctgctctcgttctcgacagcttcggggaact 65
        ||| || ||| | ||| || ||| |||||| || | ||| ||| |||
Db  933 TTACAAGGTGCCCAGCATCCACCTGCTCAAACAGGACCCTCACGTCAGCATGTTCCAGT 992

Qy   66 gattccgcagccttccaa 83
        |||||||||| |||||
Db  993 GGTCCAGGAGCAGTCCAA 1010
```

RESULT 13
ID   Q38735
     Q38735 standard; cDNA; 1723 BP.
AC   Q38735;
DT   15-JUL-1993 (first entry)
DE   Human OP-2.
KW   morphogenic; osteogenic protein; developmental cascade; hoP-2;
KW   inflammation; anti-inflammatory; Transforming Growth Factor;
KW   TGF-beta super-family; hippocampus; ss.
OS   Homo sapiens.
FH
FT   Key   Location/Qualifiers
FT   cds   490..1698
FT   /*tag= a
FT   /note= "hoP2"
FT   mat_peptide 1279..1695
FT   /*tag= b
FT   /note= "contains conserved 7 cysteine skeleton"

PN   WO9304692-A.
PD   18-MAR-1993;
PF   28-AUG-1992; U07358.
PF   30-AUG-1991; US-752764.
PR   30-AUG-1991; US-752861.
PR   30-AUG-1991; US-753059.
PA   (CREA-) CREATIVE BIOMOLECULES INC.
PI   Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E;
PI   Pang RHL, Rueger DC, Smart JE;
DR   WPI; 93-100652/12.
DR   P-PSDB; R33410.
PT   Morphogen-induced modulation of inflammatory response - and

PT resulting tissue damage, e.g. in autoimmune diseases, diabetes,
PT asthma, ischemia reperfusion injury, etc.
PS Claim 26; Page 119-121; 165pp; English.
CC Human osteogenic protein (OP)-2 is a preferred morphogen for use in
CC treating tissue damage in e.g. inflammatory disease, autoimmune
CC disease, arthritis, psoriasis, dermatitis, diabetes and emphysema.
CC Proteins having at least 70% homology with OP-2 amino acid
CC sequences can also be used. See R3400 for mature hOP-2.
SQ Sequence 1723 BP; 266 A; 625 C; 574 G; 258 T;

Query Match   29.0%; Score 25.2; DB 1; Length 1723;
Best Local Similarity   57.7%; Pred. No. 19;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   6 ttgtcagctctccatctcctctccttcagctgctctgttctcgacagcttcggggaact 65
           | ||| ||||| ||| ||||| || ||||| || ||| |||| || || |||||||| |
Db   933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCCACGTCAGCCATGTTCCAGGT 992

Qy   66 gattccgcaagccttccaa 83
           | ||| ||| ||| |||||
Db   993 GGTCCAGGAGCAGTCCAA 1010

RESULT 14
Q38859
ID Q38859 standard; DNA; 1723 BP.
AC Q38859;
DT 13-JUL-1993 (first entry)
DE Morphogen hOP2 coding sequence.
KW Morphogen; homodimer; stimulate; proliferation; progenitor cell;
KW differentiation; growth; redifferentiation; transformation; human;
KW mouse; Drosophila; Xenopus; committed cells; hippocampus; ss.
OS Homo sapiens.
FH Key   Location/Qualifiers
FT cds   490..1698
FT   /*tag= a

WPI; WO9305172-A.
PN 18-MAR-1993.
PD 28-AUG-1992; U07359.
PF 30-AUG-1991; US-752861.
PR (CREA-) CREATIVE BIOMOLECULES INC.
PA Cohen CN, Kuberasampath T, Oppermann H, Ozkaynak E, Pang RHL;
PI Rueger DC,   Smart JE;
WPI; 93-100093/12.
DR P-PSDB; R3933.
PT Screening cpds. to determine ability to modulate effective concn.
PT of a morphogen - by assaying test tissue type cells for parameter
PT indicative of a prodn. level change of morphogen
PS Disclosure; Page 95-97; 132pp; English.
CC This sequence encodes the human morphogen hOP2, isolated from the
CC hippocampus. This morphogen is inactive when reduced but is active
CC as an oxidised homodimer and when oxidised in combination with other
CC morphogens. These morphogens are capable of stimulating proliferation
CC of progenitor cell, stimulating the differentiation of progenitor
CC cells, stimulating the proliferation of differentiated cells and
CC supporting the growth and maintenance of differentiated cells,
CC including the redifferentiation of transformed cells. These
CC morphogens may also be capable of inducing redifferentiation of
CC committed cells under appropriate environmental conditions.
SQ Sequence 1723 BP; 270 A; 624 C; 572 G; 257 T;

Query Match   29.0%; Score 25.2; DB 1; Length 1723;
Best Local Similarity   57.7%; Pred. No. 19;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   6 ttgtcagctctccatctcctctccttcagctgctctgttctcgacagcttcggggaact 65
           | ||| ||||| ||| ||||| || ||||| || ||| |||| || || |||||||| |
Db   933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCCACGTCAGCCATGTTCCAGGT 992

Qy   66 gattccgcaagccttccaa 83
           | ||| ||| ||| |||||
Db   993 GGTCCAGGAGCAGTCCAA 1010

Db   993 GGTCCAGGAGCAGTCCAA 1010

RESULT 15
Q53155
ID Q53155 standard; cDNA; 1723 BP.
AC Q53155;
DT 06-JUN-1994 (first entry)
DE Sequence encoding human osteogenic protein OP2.
KW Osteogenic protein; bone; cartilage; matrix; osteoarthritis;
KW repair; vascularisation; mineralisation; differentiation; ss.
OS Homo sapiens.
FH Key   Location/Qualifiers
FT cds   490..1696
FT   /*tag= a
FT   /product= Osteogenic protein OP2.

US5266683-A.
PN 30-NOV-1993.
PD 08-APR-1988; 179406.
PF 08-APR-1988; US-179406.
PR 15-AUG-1990; US-232630.
PR 23-FEB-1989; US-315342.
PR 17-OCT-1989; US-422613.
PR 17-OCT-1989; US-422699.
PR 22-FEB-1990; US-483913.
PR 20-AUG-1990; US-569920.
PR 07-SEP-1990; US-579865.
PR 18-OCT-1990; US-599543.
PR 18-OCT-1990; US-600024.
PR 04-DEC-1990; US-621849.
PR 04-DEC-1990; US-621988.
PR 15-AUG-1988; US-236630.
PR 22-FEB-1991; US-660162.
PR 23-FEB-1989; US-560162.
PR 20-DEC-1991; US-810560.
PR 20-DEC-1991; US-810560.
PR 28-JAN-1992; US-827052.
PA 21-FEB-1992; US-841646.
PI Kuberasampath T, Oppermann H, Ozkaynak E, Pang RHL;
PI Rueger DC;
P-PSDB; R44759.
WPI; 93-395405/49.
PT New pure mammalian osteogenic proteins - induce cartilage and
PT endochondral bone formation when in association with a matrix
PS Claim 26; Columns 143-148; 128pp; English.
CC The osteogenic protein when in association with a matrix can induce
CC at the locus of an implant the full development cascade of
CC endochondral bone formation including vascularisation.
CC mineralisation and bone marrow differentiation. The osteogenic
CC protein can also be used to repair both bone and cartilage in the
CC treatment of osteoarthritis. This sequence encodes the pre-pro
CC form of the protein.
SQ Sequence 1723 BP; 265 A; 624 C; 577 G; 257 T;

Query Match   29.0%; Score 25.2; DB 1; Length 1723;
Best Local Similarity   57.7%; Pred. No. 19;
Matches   45; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   6 ttgtcagctctccatctcctctccttcagctgctctgttctcgacagcttcggggaact 65
           | ||| ||||| ||| ||||| || ||||| || ||| |||| || || |||||||| |
Db   933 TTACAAGGTGCCCAGCATCCACCTGCTCAACAGGACCCTCCACGTCAGCCATGTTCCAGGT 992

Qy   66 gattccgcaagccttccaa 83
           | ||| ||| ||| |||||
Db   993 GGTCCAGGAGCAGTCCAA 1010

Search completed: May 15, 2000, 12:08:28
Job time: 4587 sec

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 15, 2000, 11:42:02 ; Search time 1225.64 Seconds
(without alignments)
-69.052 Million cell updates/sec

Title: US-09-104-340-6
Perfect score: 87
Sequence: 1 atggattgtcagcttccat..........ttccgcagccttccaatgaa 87

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -48639572g residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_bal:*	24: em_pat:*	45: gb_htg7:*
2: gb_ba2:*	25: em_ph:*	46: em_htg1:*
3: gb_om:*	26: em_pl:*	47: em_htg2:*
4: gb_ov:*	27: em_ro:*	48: em_htg3:*
5: gb_pat:*	28: em_sts:*	49: em_hum5:*
6: gb_ph:*	29: em_sy:*	50: gb_p13:*
7: gb_pl1:*	30: em_un:*	51: gb_pr5:*
8: gb_pl2:*	31: em_vi:*	52: gb_htg8:*
9: gb_pr1:*	32: em_vl:*	53: gb_htg9:*
10: gb_pr2:*	33: gb_htg2:*	54: gb_htg10:*
11: gb_pr3:*	34: gb_in1:*	55: gb_htg11:*
12: gb_ro:*	35: gb_in2:*	56: gb_htg12:*
13: gb_sts:*	36: gb_ba1:*	57: gb_htg13:*
14: gb_sy:*	37: em_ba2:*	58: gb_htg14:*
15: gb_un:*	38: em_hum3:*	
16: gb_vi:*	39: em_hum4:*	
17: em_fun:*	40: em_pr4:*	
18: em_hum1:*	41: gb_htg3:*	
19: em_hum2:*	42: gb_htg4:*	
20: em_in:*	43: gb_htg5:*	
21: em_om:*	44: gb_htg6:*	
22: em_or:*		
23: em_ov:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		87	100.0	3132	5	A28003	A28003 H.sapiens H
2		87	100.0	3132	5	I68018	I68018 Sequence 9
3		87	100.0	3149	9	HUMHEK	HUMHEK Human recep
4		61.4	70.6	3077	12	M68515	M68515 Mouse eph-r
5		61.4	70.6	3197	12	U69278	U69278 Rattus norv
6		29	33.3	53997	44	MUSMEK4SE	MUSMEK4 Mouse eph-r
7		28.8	33.1	84551	35	RNU69278	RNU69278 Mouse eph-r
8		28.8	33.1	120625	43	AC021186	AC021186 Drosophi
9	c	28.6	32.9	119118	11	AC004295	AC004295 Drosophil
10		28.4	32.6	183776	53	AC017563	AC017563 Drosophil
11		28.4	32.6	90299	44	HSRYR7CC1	HSRYR7CC1 Human DNA
12		28.2	32.4	187701	44	AC017104	AC017104 Homo sapi
13		28.2	32.4	211385	33	HS72E17	HS72E17 Human DNA
14	c	28.2	32.4	307834	44	AL033523	AL033523 Homo sapi
15	c	28	32.2	3031	44	AC010485	AC010485 Homo sapi
16	c	28	32.2	38532	12	AL138881	AL138881 Homo sapi
17		28	32.2	74654	45	AC010475	AC010475 Homo sapi
18		28	32.2	1760	2	MMU27398	MMU27398 Mus musculu
19		27.8	32.0	5382	34	AL049628	AL049628 Streptomy
20		27.6	31.7	13095	9	AC021405	AC021405 Homo sapi
21		27.6	31.7	117270	11	AF190426	AF190426 Erwinia h
22	c	27.6	31.7	146834	45	ARDMINIC	ARDMINIC Acropora do
23	c	27.6	31.7	171950	40	Y13901	Y13901 Homo sapien
24	c	27.6	31.7	191988	43	AC005316	AC005316 Human Chr
25	c	27.6	31.7	208079	45	AC019360	AC019360 Homo sapi
26	c	27.6	31.7	113253	42	AC005317	AC005317 Human Chr
27	c	27.6	31.7	169516	43	HSFGFR4G	HSFGFR4G Homo sapien
28	c	27.4	31.5	178609	42	AC013391	AC013391 Homo sapi
29	c	27.4	31.5	179563	55	AC010121	AC010121 Drosophil
30	c	27.4	31.5	180593	55	AC013934	AC013934 Drosophil
31	c	27.4	31.5	180593	45	AC011060	AC011060 Homo sapi
32	c	27.4	31.5	213879	45	AC006433	AC006433 Homo sapi
33	c	27.4	31.5	251664	57	AC005282	AC005282 Homo sapi
34	c	27.4	31.5	319	57	AC011612	AC011612 Homo sapi
35	c	27.2	31.3	324	9	AC011607	AC011607 Homo sapi
36	c	27.2	31.3	1088	5	HUMALPHTUB	HUMALPHTUB Homo sapi
37	c	27.2	31.3	6405	9	AR051317	AR051317 Sequence
38	c	27.2	31.3	6510	11	E13168	E13168 Sequence
39	c	27.2	31.3	6740	40	AB032946	AB032946 Acacia mang
40		27.2	31.3	6990	51	AF005392	AF005392 Homo sapi
41		27.2	31.3	7768	40	AF129133	AF129133 Homo sapi
42		27.2	31.3	7898	40	AF142567	AF142567 Homo sapi
43		27.2	31.3	31756	11	AF051946	AF051946 Homo sapi
44		27.2	31.3	45896	11	Z81002	Z81002 Human DNA s
45		27.2	31.3	52097	33	AL031715	AL031715 Homo sapi

ALIGNMENTS

RESULT 1

LOCUS   A28003   3132 bp   DNA   PAT   25-SEP-1995
DEFINITION   H.saplens HEK gene.
ACCESSION   A28003
VERSION   A28003.1 GI:1247486
KEYWORDS   .
SOURCE   human.
  ORGANISM   Homo saplens
               Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
               Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3132)
  AUTHORS   .
  TITLE   A NOVEL RECEPTOR-TYPE TYROSINE KINASE AND USE THEREOF
  JOURNAL   Patent: WO 9300425-A 7 07-JAN-1993;
FEATURES   Location/Qualifiers
     source   1..3132
                     /organism="Homo saplens"
                     /db_xref="taxon:9606"
     gene   100..3051
                     /gene="HEK"
     CDS   100..3051
                     /gene="HEK"
                     /codon_start=1
                     /protein_id="CAA01906.1"
                     /db_xref="GI:1247487"
                     /db_xref="SWISS-PROT:P29320"
                     /translation="MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDIAADESFTQMDLGDR
                     ILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
                     FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKDPPSMACTRPPSSPRNVIS
                     NINETSVILDWSWPLDTGGRKDVTFNIICKKCGWNIKQCEPCSPNVRFLPRQGLTNT
                     TVTVTDLLAHTNYTFEIDAVNGVSELSSPPRQFAAVSITTNQAAPSPVLTIKKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEQETSYTILRARGTNVTISSLKPDTIYVLQ
                     IRARTAAGYGTNSRKFEFETSPDSFSISGESSQVVMIAISAAVAIILLTVVIVLIGR
                     FCGYKSKHGADERKRLHFGNGHLKLPGLRTRYVDPHTYEDPTQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLLGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDISTFR
                     TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
                     ETQSKNGPVPV"

BASE COUNT   888 a   709 c   761 g   774 t
ORIGIN

Query Match   100.0%; Score 87; DB 5; Length 3132;
Best Local Similarity   100.0%; Pred. No. 2.6e-18;
Matches   87; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 atggattgtcagctctccatcctcctcctgttcctctgttcctcgacagcttcggg 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 ATGGATTGTCAGCTCTCCATCCTCCTCCTGTTCCTCTGTTCCTCGACAGCTTCGGG 159

Qy   61 gaactgattccgcagccctccaatgaa 87
             |||||||||||||||||||||||||||
Db   160 GAACTGATTCCGCAGCCCTCCAATGAA 186

RESULT 2

LOCUS   I68018   3132 bp   DNA   PAT   04-FEB-1998
DEFINITION   Sequence 9 from patent US 5674691.
ACCESSION   I68018
VERSION   I68018.1 GI:2830140
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
               Unclassified.

Query Match   100.0%; Score 87; DB 5; Length 3132;
Best Local Similarity   100.0%; Pred. No. 2.6e-18;
Matches   87; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
ORIGIN

Qy   1 atggattgtcagctctccatcctcctcctgttcctctgttcctcgacagcttcggg 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 ATGGATTGTCAGCTCTCCATCCTCCTCCTGTTCCTCTGTTCCTCGACAGCTTCGGG 159

Qy   61 gaactgattccgcagccctccaatgaa 87
             |||||||||||||||||||||||||||
Db   160 GAACTGATTCCGCAGCCCTCCAATGAA 186

REFERENCE   1 (bases 1 to 3132)
  AUTHORS   Boyd,A.W., Simpson,R.John, Wicks,I., Ward,L.David and Wilkinson,D.
  TITLE   Method of screening for ligands to a receptor-type tyrosine kinase
  JOURNAL   Patent: US 5674691-A 9 07-OCT-1997;
FEATURES   Location/Qualifiers
     source   1..3132
                     /organism="unknown"

BASE COUNT   888 a   709 c   761 g   774 t
ORIGIN

RESULT 3

LOCUS   HUMHEK   3149 bp   mRNA   PRI   31-DEC-1994
DEFINITION   Human receptor tyrosine kinase (HEK) mRNA, complete cds.
ACCESSION   M83941
VERSION   M83941.1 GI:183931
KEYWORDS   receptor protein-tyrosine kinase.
SOURCE   Homo sapiens lymphoid tumor cDNA to mRNA.
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3149)
  AUTHORS   Wicks,I.P., Wilkinson,D., Salvaris,E. and Boyd,A.W.
  TITLE   Molecular cloning of HEK, the gene encoding a receptor tyrosine
               kinase expressed by human lymphoid tumor cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1611-1615 (1992)
  MEDLINE   92179233
FEATURES   Location/Qualifiers
     source   1..3149
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="LK63"
                     /tissue_type="lymphoid tumor"
     gene   101..3052
                     /gene="HEK"
     CDS   101..3052
                     /gene="HEK"
                     /codon_start=1
                     /product="receptor protein kinase"
                     /protein_id="AAA58633.1"
                     /db_xref="GI:183932"
                     /translation="MDCQLSILILLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDIAADESFTQMDLGDR
                     ILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
                     FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKDPPSMACTRPPSSPRNVIS
                     NINETSVILDWSWPLDTGGRKDVTFNIICKKCGWNIKQCEPCSPNVRFLPRQGLTNT
                     TVTVTDLLAHTNYTFEIDAVNGVSELSSPPRQFAAVSITTNQAAPSPVLTIKKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEQETSYTILRARGTNVTISSLKPDTIYVLQ
                     IRARTAAGYGTNSRKFEFETSPDSFSISGESSQVVMIAISAAVAIILLTVVIVLIGR
                     FCGYKSKHGADERKRLHFGNGHLKLPGLRTRYVDPHTYEDPTQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLLGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDISTFR
                     TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
                     ETQSKNGPVPV"

```
BASE COUNT          891 a      711 c      768 g      779 t
ORIGIN

Query Match
Best Local Similarity    100.0%;    Score 87;  DB 9;  Length 3149;
Matches   87;  Conservative   0;  Pred. No. 2.6e-18;

Qy      1 atggattgtcagctctccatcctcctcctttcagctgctctgttctcgaagcttcggg  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 atggattgtcagctctccatcctcctcctttcagctgctctgttctcgaagcttcggg  160

Qy     61 gaactgagtcctcagccttccaatgaa  87
          |||||||||||||||||||||||||||
Db    161 GAACTGAGTCCTCAGCCTTCCAATGAA  187
```

RESULT 4

```
LOCUS       MUSMEK4SE     2032 bp   mRNA                ROD        15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (Mek4) secreted mRNA,
            complete cds.
ACCESSION   M68515
VERSION     M68515.1  GI:454828
KEYWORDS    .
SOURCE      Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
            cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2032)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
COMMENT     On Feb 16, 1994 this sequence version replaced gi:199121.
FEATURES             Location/Qualifiers
     source          1..2032
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     CDS             89..1702
                     /codon_start=1
                     /product="Mek4 secreted"
                     /protein_id="AAA39522.1"
                     /db_xref="GI:454829"
                     /translation="MDCHLSILVLLGGCCVLSCSGELSPQPSNEVNLLDSKTIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRI
                     LKLNTEIREVGPVNKGFYLAFQDVGACVALVSRVYFKKCPFTVKNLAMFPDTVPMD
                     SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                     YKASDGAAKCAKCPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                     INETSVILDWSWPLDTGGRKDITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                     VTVTDLLAHTNYTFEIDAVNGVSELSSPPRQYAAVSITTNQAAPSPVMTIKKDRTSRN
                     SISLSWQEPEHPNGIILDYEVKYYQKEQESYTILRARGTNVTISSLKPDTTYVFQI
                     RARTAAGYGTNSRKFEFETSPDCMYFSF"
     polyA_signal    2010..2015
     polyA_site      2032
BASE COUNT          578 a      473 c      460 g      521 t
ORIGIN
```

```
Query Match
Best Local Similarity    81.6%;    Score 61.4;  DB 12;  Length 2032;
Matches   71;  Conservative   0;  Pred. No. 5.2e-10;  Mismatches 16;  Indels  0;  Gaps  0;

Qy      1 atggattgtcagctctccatcctcctcctttcagctgctctgttctcgacaagcttcggg  60
          |||||||||    |||| |||||| | ||||||||||||||||||||||||||  ||| |||
Db     89 ATGGATTGTCACCTCTCCATCCTGCTGGGCTGCTGCTGTCCTGTCGCTGCTCCTCGGGA  148
```

```
Qy     61 gaactgagtcctcagccttccaatgaa  87
          |||||| || |||||||||||||| |||
Db    149 GAACTGAGCCCACAGCCTTCCAACGAA  175
```

RESULT 5

```
LOCUS       RNU69278     3077 bp   mRNA                ROD        23-MAR-1998
DEFINITION  Rattus norvegicus eph-related receptor tyrosine kinase homolog
            (Rek4) mRNA, complete cds.
ACCESSION   U69278
VERSION     U69278.1  GI:1943913
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McTiernan,C.F. and Feldman,A.M.
  TITLE     IL-1 beta alters the expression of the receptor tyrosine kinase
            gene r-Epha3 in neonatal rat cardiomyocytes
  JOURNAL   Am. J. Physiol. 274 (1), H331-H341 (1998)
  MEDLINE   98120505
REFERENCE   2  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McTiernan,C.F. and Feldman,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1996) Cardiology, University of Pittsburgh, 200
            Lothrop Street, Pittsburgh, PA 15213, USA
COMMENT     On Apr 18, 1997 this sequence version replaced gi:1698721.
FEATURES             Location/Qualifiers
     source          1..3077
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /dev_stage="neonatal"
                     /cell_type="cardiomyocytes"
     gene            1..3077
                     /gene="Rek4"
     CDS             35..2989
                     /gene="Rek4"
                     /codon_start=1
                     /product="eph-related receptor tyrosine kinase homolog"
                     /protein_id="AAC06273.1"
                     /db_xref="GI:1698722"
                     /translation="MDCHLSILILFGCCVLSCSRELSPQPSNEVNLLDSKTIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFLEHQFTKIDTIAADESFTQMDLGDR
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPG
                     FYKALDGVAKCTKCPHSSTQEDGSMNCRCENNYFRAEKDPPSMACTRPPSAPRNVIS
                     NINETSVILDWSWPLDTGGRKDITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNT
                     TVTVTDLLAHTNYTFEIDAINGVSELSSPPRQFAAVSITLRARGTNVTISSLKPDTTYVFQ
                     IRARTAAGYGTNSRKFEFETSPDSFISGENSHVMAISAAVAIIVLTVVIVLVGR
                     FCGYHKSHSSDEKRLHFGNGHLRPLGRTIVDPHTYEDPTQAVHEFAKELDATNIAI
                     DKVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPN
                     MGVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTRGGKIPVRWTSPEAIA
                     YRKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPMDCPAALYQ
                     LMLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDIAIF
                     HTTGDWLNGMRTAHCKIFTGVEYSSCDTIAKISTDDMKKVGVTVGPQKRIISSIKA
                     LETQSKNGPVPV"
BASE COUNT          877 a      714 c      758 g      728 t
ORIGIN
```

```
Query Match
Best Local Similarity    81.6%;    Score 61.4;  DB 12;  Length 3077;
Matches   71;  Conservative   0;  Pred. No. 5.4e-10;  Mismatches 16;  Indels  0;  Gaps  0;

Qy      1 atggattgtcagctctccatcctcctcctttcagctgctctgttctcgacaagcttcggg  60
          |||||||||    |||| |||||| | ||||||||||||||||||||||||||  ||| |||
```

```
Db    35  ATGGATTGTCACCTCTCCATCCTCGTTCCGGCTCGGCTGCTGCGTCCTCCAGCTGCTCCAGG  94

Qy    61  gaactgattccgcagcctccaatgaa  87
          ||||||||||||||||||||||||||

Db    95  GAACTGAGTCCACAGCCTTCCAACGAA  121
```

```
RESULT  6

MUSMEK4

Qy    61  gaactgattccgcagcctccaatgaa  87
          |||||||||||||| |||||| ||||

Db    89  ATGGATTGTCACCTCTCCATCCTCGTTCCGGCTCGGCTGCTGCGTCCTCCAGCTGCTCCGGA  148

Qy     1  atggattgtcagctctccatcctccttcagctgctcgtttccgacagctccggg  60
          ||||||||||| ||||||||||||| || |||| || ||| ||| || ||||| |

Qy    61  gaactgattccgcagcctccaatgaa  87
          ||||||||||||||||||||||||||

Db    89  ATGGATTGTCACCTCTCCATCCTCGTTCCGGCTCGGCTGCTGCGTCCTCCAGCTGCTCCGGA  148

Qy    61  gaactgattccgcagcctccaatgaa  87
```

```
LOCUS       MUSMEK4        3197 bp    mRNA            ROD       15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (Mek4) mRNA, complete
            cds.
ACCESSION   M68513
VERSION     M68513.1  GI:199119
KEYWORDS    receptor tyrosine kinase.
SOURCE      Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo
            cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3197)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
FEATURES             Location/Qualifiers
     source          1..3197
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     gene            89..3040
                     /gene="Mek4"
     CDS             89..3040
                     /gene="Mek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA39521.1"
                     /db_xref="GI:199120"
                     /translation="MDCHLSILVLLGGCVLSCSGELSPQPSNEVNLLDSKIIQGELGW
                     ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDRSQNNWLRTNWVPRNSAQKIIVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRI
                     LKINTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVRNLAMFDTVPMD
                     SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                     YKASDAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPSMACRAPPSAPRNVISN
                     INETSVILDWSWPLDTGGRKDITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                     VTVTDLLAHTNYTFEIDAVNGVSELSSPPRQYAAVSITTNQAAPSPVMTIKDRTSRN
                     SISLSWQEPEHPNGIILLDYEYKYYQKEQEYSTILRARGRNNVTISSLKPDTIYVFQI
                     RARTAAGYGTNSRKFEFETSPDSFSIGENSHVVMIAISAAVAIVILVTVVLVLGRF
                     CGYHKSKHSAEEKRLHFGNGHLHLPGLRTYVDPHTYEDPTQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLRLPSKKEISVAIKTLKVGYTEKQRDFLGEASIMGQFDHNI
                     IRLEGVVTKSKPMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAMSY
                     RKFTSASDVWSYGIVIWEMVSYGERPYSQMSNQDVIKAVDERYRLPPPMDGPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDIATFH
                     TTGDWLNGMRTAHCKEIFTGVEYSSCDFIAKISTDDMKKVGVTVVGPQKKIISTIKAL
                     ETQSKNGPVPV"
BASE COUNT      907 a    767 c    775 g    748 t
ORIGIN
```

```
Query Match             70.6%;  Score 61.4;  DB 12;  Length 3197;
Best Local Similarity   81.6%;  Pred. No. 5.5e-10;
Matches   71;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;
```

```
Db   149  GAACTGAGCCCACAGCCTTCCAACGAA  175
          ||||||| || ||||||||||||| |||
```

```
RESULT  7

LOCUS       AC021186       53997 bp    DNA                  HTG       14-JAN-2000
DEFINITION  Homo sapiens clone RP11-744D14, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC021186
VERSION     AC021186.1  GI:6693396
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 53997)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 53997)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     Center project name: NH0563P16.
            * NOTE: This record contains 87 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
                  1:      389: contig of 389 bp in length
                390:      399: gap of unknown length
                400:     1367: contig of 968 bp in length
               1368:     1377: gap of unknown length
               1378:     2389: contig of 1012 bp in length
               2390:     2399: gap of unknown length
               2400:     3353: contig of 954 bp in length
               3354:     3363: gap of unknown length
               3364:     4194: contig of 831 bp in length
               4195:     4204: gap of unknown length
               4205:     4674: contig of 470 bp in length
               4675:     4684: gap of unknown length
               4685:     5046: contig of 362 bp in length
               5047:     5056: gap of unknown length
               5057:     5483: contig of 427 bp in length
               5484:     5493: gap of unknown length
               5494:     5647: contig of 154 bp in length
               5648:     5657: gap of unknown length
               5658:     6112: contig of 455 bp in length
               6113:     6122: gap of unknown length
               6123:     6972: contig of 850 bp in length
               6973:     6982: gap of unknown length
               6983:     7421: contig of 439 bp in length
               7422:     7431: gap of unknown length
               7432:     7835: contig of 404 bp in length
               7836:     7845: gap of unknown length
               7846:     8687: contig of 842 bp in length
               8688:     8697: gap of unknown length
               8698:     9143: contig of 446 bp in length
               9144:     9153: gap of unknown length
               9154:     9623: contig of 470 bp in length
               9624:     9633: gap of unknown length
               9634:     9766: contig of 133 bp in length
               9767:     9776: gap of unknown length
               9777:    10679: contig of 903 bp in length
              10680:    10689: gap of unknown length
```

```
10690 11095: contig of 406 bp in length
11096 11105: gap of unknown length
11106 11560: contig of 455 bp in length
11561 11570: gap of unknown length
11571 12357: contig of 787 bp in length
12358 12367: gap of unknown length
12368 12803: contig of 436 bp in length
12804 12813: gap of unknown length
12814 13287: contig of 474 bp in length
13288 13297: gap of unknown length
13298 13758: contig of 461 bp in length
13759 13768: gap of unknown length
13769 14181: contig of 413 bp in length
14182 14191: gap of unknown length
14192 14651: contig of 460 bp in length
14652 14661: gap of unknown length
14662 14986: contig of 325 bp in length
14987 14996: gap of unknown length
14997 15462: contig of 466 bp in length
15463 15472: gap of unknown length
15473 15947: contig of 475 bp in length
15948 15957: gap of unknown length
15958 16378: contig of 421 bp in length
16379 16388: gap of unknown length
16389 16857: contig of 469 bp in length
16858 16867: gap of unknown length
16868 17320: contig of 453 bp in length
17321 17330: gap of unknown length
17331 18171: contig of 841 bp in length
18172 18181: gap of unknown length
18182 19048: contig of 867 bp in length
19049 19058: gap of unknown length
19059 19973: contig of 915 bp in length
19974 19983: gap of unknown length
19984 20865: contig of 882 bp in length
20866 20875: gap of unknown length
20876 21718: contig of 843 bp in length
21719 21728: gap of unknown length
21729 22156: contig of 428 bp in length
22157 22166: gap of unknown length
22167 22624: contig of 458 bp in length
22625 22634: gap of unknown length
22635 23571: contig of 937 bp in length
23572 23581: gap of unknown length
23582 24490: contig of 909 bp in length
24491 24500: gap of unknown length
24501 25316: contig of 816 bp in length
25317 25326: gap of unknown length
25327 26304: contig of 978 bp in length
26305 26314: gap of unknown length
26315 27085: contig of 771 bp in length
27086 27095: gap of unknown length
27096 27933: contig of 838 bp in length
27934 27943: gap of unknown length
27944 28199: contig of 256 bp in length
28200 28209: gap of unknown length
28210 28982: contig of 773 bp in length
28983 28992: gap of unknown length
28993 29273: contig of 281 bp in length
29274 29283: gap of unknown length
29284 29629: contig of 346 bp in length
29630 29639: gap of unknown length
29640 29826: contig of 187 bp in length
29827 29836: gap of unknown length
29837 30588: contig of 752 bp in length
30589 30598: gap of unknown length
30599 31194: contig of 596 bp in length
31195 31204: gap of unknown length
31205 31597: contig of 393 bp in length
31598 31607: gap of unknown length
31608 32529: contig of 922 bp in length
32530 32539: gap of unknown length
32540 33475: contig of 936 bp in length

33476 33485: gap of unknown length
33486 34146: contig of 661 bp in length
34147 34156: gap of unknown length
34157 35232: contig of 1076 bp in length
35233 35242: gap of unknown length
35243 35494: contig of 252 bp in length
35495 35504: gap of unknown length
35505 36453: contig of 949 bp in length
36454 36463: gap of unknown length
36464 37079: contig of 616 bp in length
37080 37089: gap of unknown length
37090 37907: contig of 818 bp in length
37908 37917: gap of unknown length
37918 38197: contig of 280 bp in length
38198 38207: gap of unknown length
38208 38528: contig of 321 bp in length
38529 38538: gap of unknown length
38539 39502: contig of 964 bp in length
39503 39512: gap of unknown length
39513 40424: contig of 912 bp in length
40425 40434: gap of unknown length
40435 40649: contig of 215 bp in length
40650 40659: gap of unknown length
40660 41601: contig of 942 bp in length
41602 41611: gap of unknown length
41612 42577: contig of 966 bp in length
42578 42587: gap of unknown length
42588 43372: contig of 785 bp in length
43373 43382: gap of unknown length
43383 44123: contig of 741 bp in length
44124 44133: gap of unknown length
44134 44428: contig of 295 bp in length
44429 44438: gap of unknown length
44439 45402: contig of 964 bp in length
45403 45412: gap of unknown length
45413 45580: contig of 168 bp in length
45581 45590: gap of unknown length
45591 45847: contig of 247 bp in length
45848 46006: contig of 159 bp in length
46007 46016: gap of unknown length
46017 46221: contig of 205 bp in length
46222 46231: gap of unknown length
46232 47280: contig of 1049 bp in length
47281 47290: gap of unknown length
47291 48030: contig of 740 bp in length
48031 48040: gap of unknown length
48041 49181: contig of 1141 bp in length
49182 49191: gap of unknown length
49192 49814: contig of 623 bp in length
49815 49824: gap of unknown length
49825 50733: contig of 909 bp in length
50734 50743: gap of unknown length
50744 51515: contig of 772 bp in length
51516 51525: gap of unknown length
51526 51657: contig of 132 bp in length
51658 51667: gap of unknown length
51668 52726: contig of 1059 bp in length
52727 52736: gap of unknown length
52737 52899: contig of 163 bp in length
```

```
Query Match     33.3%;  Score 29;  DB 44;  Length 53997;
Best Local Similarity   63.8%;  Pred. No. 24;
Matches   44;  Conservative   0;  Mismatches   25;  Indels   0;  Gaps   0;

Qy     13 ctctccatcctcctcagctgctctcgtcctcgacagcttcggggaactgattccg 72
          ||||||| ||| |||||||| || | | |||| ||| || |||||| |||||||
Db   6898 CTCCCCCTCCCGCGCCCCCCGCCCGGCTCCCGCGCCTCCACTACGGCGCGCACACATTCCG 6957

Qy     73 cagcctcc 81
          ||| || ||
Db   6958 CCGCCAGCC 6966
```

RESULT 8
LOCUS   AC004295   84551 bp   DNA   INV   29-JUL-1998
DEFINITION Drosophila melanogaster DNA sequence (P1 DS08374 (D180)), complete
            sequence.
ACCESSION   AC004295 AC003461 AC003462 AC003463 AC003824 AC003825
            AC003465 AC003466 AC003467 AC003468 AC003469 AC003470
VERSION   AC004295.1 GI:3347818
KEYWORDS   HTG.
SOURCE   Drosophila melanogaster DNA.
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 84551)
  AUTHORS   Celniker,S.E., George,R.A., Galle,R.F., Hoskins,R.A.,
            Svirskas,R.R., Harris,N.L., Agbayani,A., Arcaina,T.T., Baxter,E.,
            Blazej,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E.,
            Flanagan,J., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,S.H., Lee,B., Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S.,
            Moshrefi,A.R., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitelaw,K.R.,
            Yee,A., Zhang,R., Zieran,L.L. and Kimmel,B.E.
  TITLE   Sequencing of Drosophila chromosome 2R, region 55C1-55C4
  JOURNAL   Unpublished (1997)
REFERENCE   2 (bases 1 to 84551)
  AUTHORS   Celniker,S.E., George,R.A., Galle,R., Svirskas,R.R., Hoskins,R.A.,
            Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Chavez,C.,
            Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
            Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
            Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R.,
            Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E.,
            Snir,E., Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R.,
            Zieran,L.L. and Kimmel,B.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-MAR-1998) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT   On Jul 29, 1998 this sequence version replaced gi:3334971.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://fruitfly.berkeley.edu/sequence/) or send
            email to drosophila@mhgc.lbl.gov.
FEATURES   Location/Qualifiers
     source   1..84551
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
                     /map="55C1-55C4"
                     /clone="P1 DS08374 (D180)"
BASE COUNT   22625 a 19310 c 19295 g 23321 t
ORIGIN

Query Match   33.1%; Score 28.8; DB 35; Length 84551;
Best Local Similarity 65.6%; Pred. No. 30;
Matches   42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   14 tctccatcctcctcagctgtctggggaactgattccgc 73
            ||||||||||||| |||||||| |||||||| | |||
Db   6538 TCTTCCTCCCTCTCCCGGTCCCGGAAACCCACTCCGG 6597

Qy   74 agcc 77
            | |
Db   6598 CGGC 6601

RESULT 9
LOCUS   AC017563   120625 bp   DNA   HTG   10-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC017563
VERSION   AC017563.1 GI:6554434
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE   fruit fly.
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 120625)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT   This sequence was identified as CDM:10211270 by the submitter.
            For more information on this record e-mail to flye@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES   Location/Qualifiers
     source   1..120625
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT   33321 a 27073 c 27008 g 33223 t
ORIGIN

Query Match   33.1%; Score 28.8; DB 43; Length 120625;
Best Local Similarity 65.6%; Pred. No. 31;
Matches   42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   14 tctccatcctcctcagctgtctggggaactgattccgc 73
            ||||||||||||| |||||||| |||||||| | |||
Db   100444 TCTTCCTCCCTCTCCCGGTCCCGGAAACCCACTCCGG 100503

Qy   74 agcc 77
            | |
Db   100504 CGGC 100507

RESULT 10
LOCUS   HSRYR7CC1   119118 bp   DNA   PRI   12-DEC-1999
DEFINITION Human DNA sequence from clone XX-PRYR7CC1 on chromosome 22 Contains
            an STS, GSSs, genomic marker D22S928, tc and ca repeat
            polymorphisms and a putative CpG island, complete sequence.
ACCESSION   AL049760
VERSION   AL049760.26 GI:5777587
KEYWORDS   HTG; ca repeat polymorphism; CpG island; D22S928; tc repeat
            polymorphism.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 119118)
  AUTHORS   Ramsay,H.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT   On Aug 26, 1999 this sequence version replaced gi:5763790.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            This sequence was generated from part of bacterial clone contigs of

```
human chromosome 22, constructed by the Sanger Centre Chromosome 22
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr22
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep IMPORTANT: This
sequence is not the entire insert of clone XX-PRYR7CCl It may be
shorter because we only sequence overlapping sections once, or
longer because we arrange for a small overlap between neighbouring
submissions.
The true left end of clone XX-PRYR7CCl is at 1 in this sequence.
The true left end of clone CTA-217C2 is at 119019 in this sequence.
The true right end of clone RP4-753M9 is at 50597 in this sequence.

FEATURES             Location/Qualifiers
     source          1..119118
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /clone="XX-PRYR7CCl"
     misc_feature    12..3186
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   77..504
                     /note="214 copies 2 mer gg 55 conserved"
     repeat_region   111..500
                     /note="15 copies 26 mer 56 conserved"
     repeat_region   369..488
                     /note="3 copies 40 mer 76 conserved"
     repeat_region   1585..1620
                     /note="6 copies 6 mer gtgagt 86 conserved"
     repeat_region   2485..2644
                     /note="4 copies 40 mer 70 conserved"
     repeat_region   2527..2598
                     /note="12 copies 6 mer ccccg 69 conserved"
     repeat_region   2593..2608
                     /note="Single clone region"
     misc_feature    2647..2669
                     /note="Single clone region"
     misc_feature    2719..2726
                     /note="Single clone region"
     misc_feature    2744..2817
                     /note="Single clone region"
     repeat_region   3678..3707
                     /note="37 copies 2 mer cc 66 conserved"
     repeat_region   3956..4249
                     /note="15 copies 2 mer tt 90 conserved"
     repeat_region   4418..4542
                     /note="AluSg repeat: matches 1..307 of consensus"
     repeat_region   5264..5329
                     /note="AluJo/FRAM repeat: matches 152..276 of consensus"
     repeat_region   5999..6097
                     /note="MIR repeat: matches 70..172 of consensus"
     repeat_region   6121..6177
                     /note="MER94 repeat: matches 47..107 of consensus"
     repeat_region   6178..6472
                     /note="AluSq repeat: matches 1..295 of consensus"
     repeat_region   6473..6500
                     /note="MER94 repeat: matches 27..47 of consensus"
     repeat_region   6501..6797
                     /note="AluY repeat: matches 1..297 of consensus"
     repeat_region   7334..7400
                     /note="MER21B repeat: matches 706..789 of consensus"
     repeat_region   7779..7894
                     /note="MER2 repeat: matches 18..132 of consensus"
     repeat_region   7895..7975

     repeat_region   8005..8080
                     /note="AluY repeat: matches 2..82 of consensus"
     repeat_region   8105..8229
                     /note="MER2 repeat: matches 269..343 of consensus"
     repeat_region   8257..8415
                     /note="AluSg/x repeat: matches 9..133 of consensus"
     repeat_region   8630..8919
                     /note="AluJo/FRAM repeat: matches 149..304 of consensus"
     repeat_region   9266..9580
                     /note="AluJo repeat: matches 13..300 of consensus"
     repeat_region   9714..9845
                     /note="AluSp repeat: matches 5..262 of consensus"
     repeat_region   9973..10170
                     /note="MIR repeat: matches 105..240 of consensus"
     repeat_region   10270..10480
                     /note="AluJo/FRAM repeat: matches 34..240 of consensus"
     repeat_region   10481..10735
                     /note="L1MA10 repeat: matches 6135..6322 of consensus"
     repeat_region   10736..10821
                     /note="AluSq repeat: matches 57..311 of consensus"
     repeat_region   10871..11171
                     /note="L1MA10 repeat: matches 6045..6135 of consensus"
     repeat_region   11190..11487
                     /note="AluJo repeat: matches 1..303 of consensus"
     repeat_region   11610..11726
                     /note="AluSx repeat: matches 1..298 of consensus"
     repeat_region   12322..12636
                     /note="MIR repeat: matches 13..134 of consensus"
     repeat_region   13016..13313
                     /note="AluSx repeat: matches 1..312 of consensus"
     repeat_region   13419..13704
                     /note="L2 repeat: matches 2616..2748 of consensus"
     repeat_region   13432..13769
                     /note="143 copies 2 mer gg 59 conserved"
     repeat_region   13878..14208
                     /note="13 copies 26 mer 57 conserved"
     repeat_region   14224..14886
                     /note="MIR1C repeat: matches 6..368 of consensus"
     repeat_region   14893..15204
                     /note="L1PA10 repeat: matches 5476..6165 of consensus"
     repeat_region   15226..15335
                     /note="AluSp repeat: matches 1..312 of consensus"
     repeat_region   15336..15446
                     /note="L1P3 repeat: matches 5356..5465 of consensus"
     repeat_region   15484..15777
                     /note="MIT1C repeat: matches 355..464 of consensus"
     repeat_region   complement(15768..16376)
                     /note="AluSq repeat: matches 1..295 of consensus"
     misc_feature    16387..16541
                     /note="match: GSS: Em:AQ476535"
     repeat_region   17071..17122
                     /note="MIR repeat: matches 73..226 of consensus"
     repeat_region   17222..17308
                     /note="26 copies 2 mer aa 73 conserved"
     repeat_region   17452..17487
                     /note="L2 repeat: matches 2652..2738 of consensus"
     repeat_region   17599..17908
                     /note="18 copies 2 mer tg 91 conserved"
     repeat_region   18508..18616
                     /note="AluY repeat: matches 1..310 of consensus"
     repeat_region   18668..18965
                     /note="L1PA6 repeat: matches 6035..6143 of consensus"
     repeat_region   19069..19100
                     /note="AluSx repeat: matches 1..298 of consensus"
     repeat_region   19494..19849
                     /note="MLT1-INTERNAL repeat: matches 692..723 of consensus"
     repeat_region   19869..20149
                     /note="MLT1-INTERNAL repeat: matches 927..1296 of consensus"
     repeat_region   20218..20385
                     /note="AluJo repeat: matches 18..288 of consensus"
     repeat_region   ...20385
                     /note="MLT1B repeat: matches 1..180 of consensus"
```

```
repeat_region   20386..20684
                /note="AluSp repeat: matches 1..303 of consensus"
repeat_region   20685..20736
                /note="MlT1B repeat: matches 180..234 of consensus"
repeat_region   20737..21016
                /note="AluSg repeat: matches 1..280 of consensus"
repeat_region   21017..21215
                /note="MlT1B repeat: matches 234..390 of consensus"
repeat_region   21217..21347
                /note="MIR repeat: matches 12..144 of consensus"
repeat_region   21582..21731
                /note="L2 repeat: matches 2511..2657 of consensus"
repeat_region   22386..22415
                /note="15 copies 2 mer ac 93 conserved"
repeat_region   23579..23881
                /note="AluSx repeat: matches 1..310 of consensus"
misc_feature    complement(23985..24510)
                /note="match: GSS: Em:AQ529159"
misc_feature    complement(24086..24604)
                /note="match: GSS: Em:AQ569063"
repeat_region   24115..24201
                /note="MlT1D repeat: matches 6..94 of consensus"
repeat_region   24201..24304
                /note="MlT1B repeat: matches 362..466 of consensus"
misc_feature    complement(24215..24582)
                /note="match: GSS: Em:AQ375812"
misc_feature    24487..24919
                /note="match: GSS: Em:AQ206604"

Query Match
Best Local Similarity   32.9%;
Matches  43; Conservative  0; Pred. No. 35;   DB 11;  Length 119118;
                                                 Score 28.6;
                                 Mismatches  24; Indels  0;  Gaps  0;

Qy     13 ctctccatcctcctcctcttcagctgcctctgttcctgggggggaactgattccg 72
          ||||-|||-|||-||||||-||-||||||-||-|||||||-||-|||-||-|-|||
Db  12792 CGCCCCTGCGCGTGCTGCTTCTGAAATCTCCCCTGGCTCTCGAACCTCTCCGGGGTTGTGATGCGG 12851

Qy     73 cagcctt 79
          ||||||
Db  12852 CCCCCTT 12858

RESULT  11
LOCUS       AC017104   183776 bp  DNA   HTG   02-FEB-2000
DEFINITION  Homo sapiens clone RP11-56215, WORKING DRAFT SEQUENCE, 25 unordered
            pieces.
ACCESSION   AC017104
VERSION     AC017104.3  GI:6852250
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 183776)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 183776)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 2, 2000 this sequence version replaced gi:6693437.
            -------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information -----------------
            Center project name: H_NH0562105
            -------------- Summary Statistics -----------------
```

```
            *  Sequencing vector: M13; 88%
            *  Sequencing vector: plasmid; 12%
            *  Chemistry: Dye-primer ET; 88% of reads
            *  Chemistry: Dye-terminator Big Dye; 12% of reads
            *  Assembly program: Phrap; version 0.990319
            *  Consensus quality: 168893 bases at least Q40
            *  Consensus quality: 174250 bases at least Q30
            *  Consensus quality: 177018 bases at least Q20
            *  Insert size: 173000; agarose-fp
            *  Insert size: 183776; sum-of-contigs
            *  Quality coverage: 3.77 in Q20 bases; agarose-fp
            *  Quality coverage: 3.55 in Q20 bases; sum-of-contigs
            *
            *  NOTE: This is a 'working draft' sequence. It currently
            *  consists of 25 contigs. The true order of the pieces
            *  is not known and their order in this sequence record is
            *  arbitrary. Gaps between the contigs are represented as
            *  runs of N, but the exact sizes of the gaps are unknown.
            *  This record will be updated with the finished sequence
            *  as soon as it is available and the accession number will
            *  be preserved.
            *
            *      1       1219:  contig of 1219 bp in length
            *   1220       3630:  gap of unknown length
            *   3631       5029:  contig of 1399 bp in length
            *   5030       8124:  gap of unknown length
            *   8125       9784:  contig of 1660 bp in length
            *   9785      13985:  gap of unknown length
            *  13986      16195:  contig of 2210 bp in length
            *  16196      19858:  gap of unknown length
            *  19859      24742:  contig of 4884 bp in length
            *  24743      31343:  gap of unknown length
            *  31344      36204:  contig of 4861 bp in length
            *  36205      41101:  gap of unknown length
            *  41102      47115:  contig of 6014 bp in length
            *  47116      54793:  gap of unknown length
            *  54794      62416:  contig of 7623 bp in length
            *  62417      70405:  gap of unknown length
            *  70406      76869:  contig of 6464 bp in length
            *  76870      85404:  gap of unknown length
            *  85405      94187:  contig of 8783 bp in length
            *  94188     103507:  gap of unknown length
            * 103608     113100:  contig of 9493 bp in length
            * 113101     127447:  gap of unknown length
            * 127448     141478:  contig of 14031 bp in length
            * 141479     158627:  gap of unknown length
            * 158628     183776:  contig of 25149 bp in length.

FEATURES             Location/Qualifiers
     source          1..183776
                     /organism="Homo sapiens"
```

BASE COUNT   46803 a   44968 c   44823 g   47139 t   43 others
ORIGIN

/db_xref="taxon:9606"
/clone="RP11-562I5"

RESULT 12
HS72E17
Query Match   32.68; Score 28.4; DB 53; Length 183776;
Best Local Similarity 62.9%; Pred. No. 43;
Matches   44; Conservative   0; Mismatches   26; Indels   0; Gaps   0;

Qy   2 tggattgtcagctctccatccctccttccagctcgctcgttctcgacagcttcgggg 61
          ||||||||||||||||||||||||||||||||||||||||||| |||| || ||||
Db 12751 TGCAANGCAACCCTCCAGCCTCTCGCCCCCAGAGAGGAGCACTGTTGTCAACAGCAGCCTGA 12692

Qy   62 aactgattcc 71
          |||||| ||||
Db 12691 GGCTGAATCC 12682

LOCUS   HS72E17   90299 bp   DNA   PRI   29-NOV-1999
DEFINITION Human DNA sequence from clone RP1-72E17 on chromosome 6p24.1-25.3,
            complete sequence.
ACCESSION   AL033523
VERSION   AL033523.14 GI:6478156
KEYWORDS   HTG.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 90299)
AUTHORS   Milne,S.
TITLE   Direct Submission
JOURNAL   Submitted (29-NOV-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequests@sanger.ac.uk
COMMENT   On Nov 29, 1999 this sequence version replaced gi:6468554.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-72E17 is from the library RPCI-1 constructed at the Roswell
            Park Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP1-72E17.
FEATURES   Location/Qualifiers
     source   1..90299
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p24.1-25.3"
                     /clone="RP1-72E17"
                     /clone_lib="RPCI-1"
     misc_feature   7992..8401

misc_feature   /note="match: GSS: Em:AQ342290.1; match: STS: Em:G58015.1"
                8184..8478
misc_feature   /note="match: GSS: Em:AQ101900"
                11636..11917
misc_feature   /note="match: STS: Em:F05237"
                17034..17305
                /note="match: GSS: Em:AQ122350"
misc_feature   complement(12136..12380)
                /note="match: STS: Em:F01490"
misc_feature   complement(17175..17672)
                /note="match: GSS: Em:G28114; match: STS: Em:T91135"
misc_feature   complement(17272..17672)
                /note="match: STS: Em:G25929"
misc_feature   complement(41270..41845)
                /note="match: STS: Em:AA131878"
misc_feature   complement(41285..41846)
                /note="match: STS: Em:G23248"
misc_feature   41498..41706
                /note="match: STS: Em:T90903"
misc_feature   complement(46466..46874)
                /note="match: STS: Em:AA437259"
misc_feature   complement(50974..51452)
                /note="match: STS: Em:AQ587947.1"
misc_feature   53812..54273
                /note="match: GSS: Em:AQ223851"
misc_feature   complement(54830..55149)
                /note="match: GSS: Em:AQ039120"
misc_feature   55156..55504
                /note="match: GSS: Em:AQ022131"
misc_feature   55156..55631
                /note="match: GSS: Em:B93140"
misc_feature   55156..55566
                /note="match: GSS: Em:AQ125812"
misc_feature   69199..69703
                /note="match: GSS: Em:AQ118077"
misc_feature   complement(70389..70829)
                /note="match: STS: Em:AA774466"
misc_feature   complement(74369..74765)
                /note="match: GSS: Em:AA774065"
misc_feature   80516..81048
                /note="match: GSS: Em:AQ184450"
misc_feature   80535..81048
                /note="match: GSS: Em:AQ767778.1"

BASE COUNT   24639 a   20192 c   20337 g   25131 t
ORIGIN

Query Match   32.48; Score 28.2; DB 11; Length 90299;
Best Local Similarity 59.3%; Pred. No. 47;
Matches   48; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   7 tgtcaagctctccatccctccttccagctcgctcgttctcgacagcttcggggaactg 66
          || ||||| ||||||||||||||||||||||| ||| || ||| || |||| ||| |
Db 75509 TGCCAGCAACCTCTTACTCCATCCCACTCTCGCCCTCGCCCCCTGATCCCAACATAGTCCTGGGAATTT 75568

Qy   67 atccgcagcctccaatgaa 87
          |||| || || ||||| ||
Db 75569 TTTCCCATGCAGTTGAATCAA 75589

RESULT 13
AC010485/c
LOCUS   AC010485   187701 bp   DNA   HTG   14-JAN-2000
DEFINITION Homo sapiens chromosome 19 clone CITB-H1_2329C7, WORKING DRAFT
            SEQUENCE, 44 unordered pieces.
ACCESSION   AC010485
VERSION   AC010485.2 GI:6693194
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 187701)
AUTHORS   DOE Joint Genome Institute.
TITLE   Sequencing of Human Chromosome 19

```
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 187701)
AUTHORS      DOE Joint Genome Institute.
TITLE        Direct Submission
JOURNAL      Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
             Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT      On Jan 14, 2000 this sequence version replaced gi:5882452.
             -----Genome Center
               Center: Joint Genome Institute
               Center Code: JGI
               Web site: http://www.jgi.doe.gov
             -----Summary Statistics
               Consensus quality: 129296 bases at least Q40
               Consensus quality: 142143 bases at least Q30
               Consensus quality: 147093 bases at least Q20
               Consensus quality:  187701; sum-of-contigs estimation
               Estimated insert size: 125880; agarose-fp estimation
               Estimated insert size: 187701; sum-of-contigs estimation
               Quality coverage:  8.86x in Q20 bases; agarose-fp estimation
               Quality coverage:  5.95x in Q20 bases; sum-of-contigs estimation
             -----.
             *
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 44 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *
             *   1     1460: contig of 1460 bp in length
             *            1460: gap of unknown length
             *   1461     2605: contig of 1145 bp in length
             *            2605: gap of unknown length
             *   2606     4240: contig of 1635 bp in length
             *            4240: gap of unknown length
             *   4241     5352: contig of 1112 bp in length
             *            5352: gap of unknown length
             *   5353     6469: contig of 1117 bp in length
             *            6469: gap of unknown length
             *   6470     8438: contig of 1969 bp in length
             *            8438: gap of unknown length
             *   8439    10465: contig of 2027 bp in length
             *           10465: gap of unknown length
             *  10466    11680: contig of 1215 bp in length
             *           11680: gap of unknown length
             *  11681    13478: contig of 1798 bp in length
             *           13478: gap of unknown length
             *  13479    15562: contig of 2084 bp in length
             *           15562: gap of unknown length
             *  15563    16848: contig of 1286 bp in length
             *           16848: gap of unknown length
             *  16849    18145: contig of 1297 bp in length
             *           18145: gap of unknown length
             *  18146    19498: contig of 1353 bp in length
             *           19498: gap of unknown length
             *  19499    20634: contig of 1136 bp in length
             *           20634: gap of unknown length
             *  20635    22236: contig of 1602 bp in length
             *           22236: gap of unknown length
             *  22237    23612: contig of 1376 bp in length
             *           23612: gap of unknown length
             *  23613    24794: contig of 1182 bp in length
             *           24794: gap of unknown length
             *  24795    27238: contig of 2444 bp in length
             *           27238: gap of unknown length
             *  27239    28627: contig of 1389 bp in length
             *           28627: gap of unknown length
             *  28628    30179: contig of 1552 bp in length
             *           30179: gap of unknown length
             *  30180    31616: contig of 1437 bp in length
             *           31616: gap of unknown length
             *  31617    32928: contig of 1312 bp in length
             *           32928: gap of unknown length
             *  32929    34098: contig of 1170 bp in length
```

```
             *  34099    35294: contig of 1196 bp in length
             *           35294: gap of unknown length
             *  35295    36809: contig of 1515 bp in length
             *           36809: gap of unknown length
             *  36810    38428: contig of 1619 bp in length
             *           38428: gap of unknown length
             *  38429    39806: contig of 1378 bp in length
             *           39806: gap of unknown length
             *  39807    41918: contig of 2112 bp in length
             *           41918: gap of unknown length
             *  41919    43164: contig of 1246 bp in length
             *           43164: gap of unknown length
             *  43165    44462: contig of 1298 bp in length
             *           44462: gap of unknown length
             *  44463    45613: contig of 1151 bp in length
             *           45613: gap of unknown length
             *  45614    47443: contig of 1830 bp in length
             *           47443: gap of unknown length
             *  47444    48948: contig of 1505 bp in length
             *           48948: gap of unknown length
             *  48949    50801: contig of 1853 bp in length
             *           50801: gap of unknown length
             *  50802    52829: contig of 2028 bp in length
             *           52829: gap of unknown length
             *  52830    54515: contig of 1686 bp in length
             *           54515: gap of unknown length
             *  54516    56458: contig of 1943 bp in length
             *           56458: gap of unknown length
             *  56459    59472: contig of 3014 bp in length
             *           59472: gap of unknown length
             *  59473    60769: contig of 1297 bp in length
             *           60769: gap of unknown length
             *  60770    69868: contig of 9099 bp in length
             *           69868: gap of unknown length
             *  69869    83413: contig of 13545 bp in length
             *           83413: gap of unknown length
             *  83414   102397: contig of 18984 bp in length
             *          102397: gap of unknown length
             * 102398   124845: contig of 22448 bp in length
             *          124845: gap of unknown length
             * 124846   187701: contig of 62856 bp in length.
             *          187701: gap of unknown length.
FEATURES             Location/Qualifiers
     source          1..187701
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CITB-H1_2329C7"
BASE COUNT     47607 a   47529 c   45578 g   46667 t     320 others
ORIGIN
```

```
Qy       8 gtcagctctccatcctcctcctccttcagctgctctgttctcgacaagcttcgggggaactga 67
           | |||||    |||||||||||  |||||  |||||| || ||||||||||||
Db  181274 GACAGCTTCCCCTCTCCCTCCTTCCAGCGCTGCTTCCATACTAGATTCGGGGCCATTT 181215

Qy      68 ttccgcagcctc 80
           | |||||| |||
Db  181214 TGCCGCAGGATTC 181202
```

```
Query Match          32.4%;  Score 28.2;  DB 44;  Length 187701;
Best Local Similarity  61.6%;  Pred. No. 50;
Matches   45;  Conservative   0;  Mismatches  28;  Indels   0;  Gaps   0;
```

```
RESULT   14
LOCUS        AL138881      211385 bp    DNA                  HTG
DEFINITION   Homo sapiens chromosome 6 clone RP11-528H16, *** SEQUENCING IN
             PROGRESS ***, 50 unordered pieces.
ACCESSION    AL138881
VERSION      AL138881.1  GI:6982856
KEYWORDS     HTG; HTGS_PHASE1.
```

SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 211385)
  AUTHORS   Sims,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (05-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT   On Feb 16, 2000 this sequence version replaced gi:6912160.
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence. Work on the sequence is in
            progress and the release of this data is based on the understanding
            that the sequence may change as work continues. The sequence may
            be contaminated with foreign sequence from E.coli, yeast, vector,
            phage etc. Order of segments is not known; 800 n's separate
            segments. Contig_ID: 00006   Length: 9572bp
                Contig_ID: 00011   Length: 1016bp
                Contig_ID: 00022   Length: 1557bp
                Contig_ID: 00032   Length: 2972bp
                Contig_ID: 00050   Length: 4358bp
                Contig_ID: 00051   Length: 5737bp
                Contig_ID: 00052   Length: 1925bp
                Contig_ID: 00082   Length: 4685bp
                Contig_ID: 00086   Length: 4360bp
                Contig_ID: 00185   Length: 2863bp
                Contig_ID: 00199   Length: 6154bp
                Contig_ID: 00206   Length: 2440bp
                Contig_ID: 00210   Length: 2803bp
                Contig_ID: 00238   Length: 1982bp
                Contig_ID: 00243   Length: 4711bp
                Contig_ID: 00263   Length: 6857bp
                Contig_ID: 00281   Length: 3004bp
                Contig_ID: 00295   Length: 3422bp
                Contig_ID: 00313   Length: 2375bp
                Contig_ID: 00337   Length: 2333bp
                Contig_ID: 00386   Length: 4205bp
                Contig_ID: 00464   Length: 1855bp
                Contig_ID: 00472   Length: 3313bp
                Contig_ID: 00473   Length: 3603bp
                Contig_ID: 00484   Length: 1719bp
                Contig_ID: 00493   Length: 4386bp
                Contig_ID: 00499   Length: 1451bp
                Contig_ID: 00506   Length: 1159bp
                Contig_ID: 00508   Length: 3841bp
                Contig_ID: 00518   Length: 1755bp
                Contig_ID: 00546   Length: 3386bp
                Contig_ID: 00549   Length: 1984bp
                Contig_ID: 00550   Length: 2332bp
                Contig_ID: 00551   Length: 3705bp
                Contig_ID: 00564   Length: 4685bp
                Contig_ID: 00580   Length: 3614bp
                Contig_ID: 00581   Length: 6142bp
                Contig_ID: 00621   Length: 4868bp
                Contig_ID: 00649   Length: 3054bp
                Contig_ID: 00651   Length: 1455bp
                Contig_ID: 00654   Length: 1855bp
                Contig_ID: 00666   Length: 2320bp
                Contig_ID: 00726   Length: 4142bp
                Contig_ID: 00737   Length: 3880bp
                Contig_ID: 00774   Length: 2731bp
                Contig_ID: 00876   Length: 352bp
                Contig_ID: 00889   Length: 950bp
                Contig_ID: 00906   Length: 2452bp
                Contig_ID: 00916   Length: 1319bp
                Contig_ID: 00955   Length: 2346bp
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 50 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence

            * as soon as it is available and the accession number will
            * be preserved.
            *   1   9572: contig of 9572 bp in length
            *   10373 10372: gap of   800 bp in length
            *   11389 11388: contig of 1016 bp in length
            *   11389 12188: gap of   800 bp in length
            *   13746 13745: contig of 1557 bp in length
            *   14546 14545: gap of   800 bp in length
            *   17518 17517: contig of 2972 bp in length
            *   18318 18317: gap of   800 bp in length
            *   22676 22675: contig of 4358 bp in length
            *   23476 23475: gap of   800 bp in length
            *   29213 29212: contig of 5737 bp in length
            *   30013 30012: gap of   800 bp in length
            *   31938 31937: contig of 1925 bp in length
            *   32738 32737: gap of   800 bp in length
            *   37423 37422: contig of 4685 bp in length
            *   38223 38222: gap of   800 bp in length
            *   42583 42582: contig of 4360 bp in length
            *   43383 43382: gap of   800 bp in length
            *   46246 46245: contig of 2863 bp in length
            *   47046 47045: gap of   800 bp in length
            *   53200 53199: contig of 6154 bp in length
            *   54000 53999: gap of   800 bp in length
            *   56440 56439: contig of 2440 bp in length
            *   57240 57239: gap of   800 bp in length
            *   60043 60042: contig of 2803 bp in length
            *   60843 60842: gap of   800 bp in length
            *   62825 62824: contig of 1982 bp in length
            *   63625 63624: gap of   800 bp in length
            *   68336 68335: contig of 4711 bp in length
            *   69136 69135: gap of   800 bp in length
            *   75993 75992: contig of 6857 bp in length
            *   76793 76792: gap of   800 bp in length
            *   79797 79796: contig of 3004 bp in length
            *   80597 80596: gap of   800 bp in length
            *   84019 84018: contig of 3422 bp in length
            *   84819 84818: gap of   800 bp in length
            *   87194 87193: contig of 2375 bp in length
            *   87994 87993: gap of   800 bp in length
            *   90327 90326: contig of 2333 bp in length
            *   91127 91126: gap of   800 bp in length
            *   95332 95331: contig of 4205 bp in length
            *   96132 96131: gap of   800 bp in length
            *   97987 97986: contig of 1855 bp in length
            *   98787 98786: gap of   800 bp in length
            *   102100 102099: contig of 3313 bp in length
            *   102900 102899: gap of   800 bp in length
            *   106503 106502: contig of 3603 bp in length
            *   107303 107302: gap of   800 bp in length
            *   109022 109021: contig of 1719 bp in length
            *   109822 109821: gap of   800 bp in length
            *   114208 114207: contig of 4386 bp in length
            *   115008 115007: gap of   800 bp in length
            *   116459 116458: contig of 1451 bp in length
            *   117259 117258: gap of   800 bp in length
            *   118418 118417: contig of 1159 bp in length
            *   119218 119217: gap of   800 bp in length
            *   123059 123058: contig of 3841 bp in length
            *   123859 123858: gap of   800 bp in length
            *   125614 125613: contig of 1755 bp in length
            *   126414 126413: gap of   800 bp in length
            *   129800 129799: contig of 3386 bp in length
            *   130600 130599: gap of   800 bp in length
            *   132584 132583: contig of 1984 bp in length
            *   133384 133383: gap of   800 bp in length
            *   135716 135715: contig of 2332 bp in length
            *   136516 136515: gap of   800 bp in length
            *   140221 140220: contig of 3705 bp in length
            *   141021 141020: gap of   800 bp in length
            *   145706 145705: contig of 4685 bp in length
            *   146506 146505: gap of   800 bp in length
            *   150119 150119: contig of 3614 bp in length

```
LOCUS       AC010475   307834 bp   DNA          HTG       14-JAN-2000
DEFINITION  Homo sapiens chromosome 5 clone CITB-H1_2308B20, WORKING DRAFT
            SEQUENCE, 59 unordered pieces.
ACCESSION   AC010475
VERSION     AC010475.2  GI:6693197
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 307834)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 307834)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
```

```
COMMENT     On Jan 14, 2000 this sequence version replaced gi:5882462.
            ------Genome Center
              Center: Joint Genome Institute
              Center Code: JGI
              Web site: http://www.jgi.doe.gov
            ------Summary Statistics
              Consensus quality: 224141 bases at least Q40
              Consensus quality: 249816 bases at least Q30
              Consensus quality: 258890 bases at least Q20
              Estimated Insert size: 307834; sum-of-contigs estimation
              Estimated Insert size: 120000; pulse field gel estimation
              Estimated Insert size: 307834; sum-of-contigs estimation
              Quality coverage: 15.47x in Q20 bases; pulse field gel estimation
              Quality coverage: 6.03x in Q20 bases; sum-of-contigs estimation
            ------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 59 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1       1120: contig of 1120 bp in length
            *              1120:      gap of unknown length
            *      1121       2849: contig of 1729 bp in length
            *              2849:      gap of unknown length
            *      2850       4226: contig of 1377 bp in length
            *              4226:      gap of unknown length
            *      4227       5364: contig of 1138 bp in length
            *              5364:      gap of unknown length
            *      5365       6490: contig of 1126 bp in length
            *              6490:      gap of unknown length
            *      6491       8454: contig of 1964 bp in length
            *              8454:      gap of unknown length
            *      8455      10829: contig of 2375 bp in length
            *             10829:      gap of unknown length
            *     10830      12773: contig of 1944 bp in length
            *             12773:      gap of unknown length
            *     12774      14208: contig of 1435 bp in length
            *             14208:      gap of unknown length
            *     14209      15373: contig of 1165 bp in length
            *             15373:      gap of unknown length
            *     15374      16618: contig of 1245 bp in length
            *             16618:      gap of unknown length
            *     16619      18294: contig of 1676 bp in length
            *             18294:      gap of unknown length
            *     18295      19533: contig of 1239 bp in length
            *             19533:      gap of unknown length
            *     19534      20661: contig of 1128 bp in length
            *             20661:      gap of unknown length
            *     20662      21910: contig of 1249 bp in length
            *             21910:      gap of unknown length
            *     21911      23186: contig of 1276 bp in length
            *             23186:      gap of unknown length
            *     23187      25653: contig of 2467 bp in length
            *             25653:      gap of unknown length
            *     25654      26815: contig of 1162 bp in length
            *             26815:      gap of unknown length
            *     26816      28446: contig of 1631 bp in length
            *             28446:      gap of unknown length
            *     28447      29673: contig of 1227 bp in length
            *             29673:      gap of unknown length
            *     29674      30852: contig of 1179 bp in length
            *             30852:      gap of unknown length
            *     30853      32175: contig of 1323 bp in length
            *             32175:      gap of unknown length
            *     32176      34081: contig of 1906 bp in length
            *             34081:      gap of unknown length
            *     34082      35819: contig of 1738 bp in length
            *             35819:      gap of unknown length
            *     35820      37630: contig of 1811 bp in length
            *             37630:      gap of unknown length
            *     37631      38840: contig of 1210 bp in length
```

```
FEATURES             Location/Qualifiers
     source          1..211385
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-528H16"
                     /clone_lib="RPCI-11.2"
BASE COUNT     47247 a   38300 c   39283 g   47285 t   39270 others
ORIGIN
     1  .211385
```

```
*   150120
*   150920  150919: gap of   6142 bp in length
*   157061  157862: gap of    800 bp in length
*   157862  162729: contig of 4868 bp in length
*   162730  163529: gap of    800 bp in length
*   163530  166583: contig of 3054 bp in length
*   166584  167383: gap of    800 bp in length
*   167384  168838: contig of 1455 bp in length
*   168839  169638: gap of    800 bp in length
*   169639  171493: contig of 1855 bp in length
*   171494  172293: gap of    800 bp in length
*   172294  174613: contig of 2320 bp in length
*   174614  175413: gap of    800 bp in length
*   175414  179555: contig of 4142 bp in length
*   179556  180355: gap of    800 bp in length
*   180356  184235: contig of 3880 bp in length
*   184236  185035: gap of    800 bp in length
*   185036  187765: contig of 2731 bp in length
*   187766  188565: gap of    800 bp in length
*   188567  192117: contig of 3552 bp in length
*   192118  192917: gap of    800 bp in length
*   192918  202867: contig of 9950 bp in length
*   202868  203667: gap of    800 bp in length
*   203668  206119: contig of 2452 bp in length
*   206120  206920: gap of    800 bp in length
*   206921  208238: contig of 1319 bp in length
*   208239  209038: gap of    800 bp in length
*   208240  209039: gap of    800 bp in length
*   209040  211385: contig of 2346 bp in length
```

```
RESULT 15
AC010475/c
LOCUS       AC010475   307834 bp   DNA          HTG       14-JAN-2000
DEFINITION  Homo sapiens chromosome 5 clone CITB-H1_2308B20, WORKING DRAFT
            SEQUENCE, 59 unordered pieces.
```

```
Query Match      32.4%;  Score 28.2;  DB 33;  Length 211385;
Best Local Similarity   59.3%;
Matches   48;  Conservative  0;  Mismatches  33;  Indels  0;  Gaps  0;
                                                       Pred. No. 50;
```

```
Qy      7 tgtcagctcctccatcctcctcctcctcagctgtctcgttctcgacagctcgggaactg 66
          ||||||| || ||||| |||||||||||||||||||||||| ||||||||||  |||
Db 205176 TGCCAGCACCTCTACTCCATCCCACTCCTCTCCCCTGATCCGACATAGTCCTGGAATTT 205235

Qy     67 attccgcagcctccaatgaa 87
          |||| || || ||| ||
Db 205236 TTTCCCATGCAGTTGAATCAA 205256
```

```
FEATURES             Location/Qualifiers
     source          1..307834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CITB-H1_2308B20"
BASE COUNT     87825 a   67908 c   67457 g   84063 t    581 others

*
*        38841    40042: gap of unknown length
*        40043    41631: contig of 1589 bp in length
*        41632    42937: gap of unknown length
*        42938    44046: contig of 1109 bp in length
*        44047    45451: gap of unknown length
*        45452    47075: contig of 1624 bp in length
*        47076    48875: gap of unknown length
*        48876    50620: contig of 1745 bp in length
*        50621    51956: gap of unknown length
*        51957    53979: contig of 2023 bp in length
*        53980    55408: gap of unknown length
*        55409    56518: contig of 1110 bp in length
*        56519    57895: gap of unknown length
*        57896    59316: contig of 1421 bp in length
*        59317    60803: gap of unknown length
*        60804    62554: contig of 1751 bp in length
*        62555    64100: gap of unknown length
*        64101    66567: contig of 2467 bp in length
*        66568    68184: gap of unknown length
*        68185    70971: contig of 2787 bp in length
*        70972    73787: gap of unknown length
*        73788    76676: contig of 2889 bp in length
*        76677    84224: gap of unknown length
*        84225    98029: contig of 13805 bp in length
*        98030   110318: gap of unknown length
*       110319   119812: contig of 9494 bp in length
*       119813   136488: gap of unknown length
*       136489   150409: contig of 13921 bp in length
*       150410   169220: gap of unknown length
*       169221   197236: contig of 28016 bp in length
*       197237   219772: gap of unknown length
*       219773   247056: contig of 27284 bp in length
*       247057   301756: gap of unknown length
*                307834: contig of 6078 bp in length.
```

```
ORIGIN
Query Match    32.4%;   Score 28.2;  DB 44;  Length 307834;
Best Local Similarity   61.6%;  Pred. No. 52;
Matches   45;  Conservative   0;  Mismatches   28;  Indels   0;  Gaps   0;

Qy          8 gtcagctctccatcctcctcagctgtctgtttctcgacagcttcggggaactga 67
              | ||||| ||| ||  ||||||| | || | | | |||  | |||||||  |
Db     297411 GACAGCTTCCCCTCTCCCTCTTCCCAGCCTGCTGCTTCCATACTAGATTCGGGCGCCATTT 297352

Qy         68 ttccgcagccttc 80
              | |||||| |||
Db     297351 TGCCGCAGGATTC 297339

Search completed: May 15, 2000, 11:43:02
Job time: 18609 sec
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 11:20:15 ; Search time 1358.94 Seconds
                           (without alignments)
                           196.854 Million cell updates/sec

Title: US-09-104-340-7
Perfect score: 66
Sequence: 1 gtcaatcactggattcaaa.........tctcttatccatcacatggg 66

Scoring table: IDENTITY_NUC
                           Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   EST:*

#	Summary	#	Summary	#	Summary
1:	em_est1:*	45:	gb_est26:*	94:	gb_gss9:*
2:	em_est2:*	46:	gb_est27:*	95:	gb_gss5:*
3:	em_est3:*	47:	gb_est28:*	96:	gb_gss6:*
4:	em_est4:*	48:	gb_est29:*	97:	em_gss7:*
5:	em_est5:*	49:	gb_est30:*	98:	em_gss8:*
6:	em_est6:*	50:	gb_est31:*	99:	em_gss9:*
7:	em_est7:*	51:	gb_est32:*	100:	em_gss10:*
8:	em_est8:*	52:	em_est20:*	101:	em_gss11:*
9:	em_est9:*	53:	em_est21:*	102:	gb_gss10:*
10:	em_est10:*	54:	em_est22:*	103:	gb_gss11:*
11:	em_est11:*	55:	em_est23:*	104:	em_gss12:*
12:	em_est12:*	56:	em_est24:*	105:	gb_gss12:*
13:	em_est13:*	57:	em_est25:*	106:	gb_gss13:*
14:	em_est14:*	58:	em_est26:*	107:	gb_gss14:*
15:	em_est15:*	59:	gb_est33:*	108:	gb_gss15:*
16:	em_est16:*	60:	gb_est34:*	109:	gb_gss16:*
17:	em_est17:*	61:	gb_est35:*		
18:	em_est18:*	62:	gb_est36:*		
19:	em_est19:*	63:	gb_est37:*		
20:	em_est1:*	64:	gb_est38:*		
21:	gb_est1:*	65:	em_est27:*		
22:	gb_est3:*	66:	em_est28:*		
23:	gb_est4:*	67:	em_est29:*		
24:	gb_est5:*	68:	em_est30:*		
25:	gb_est6:*	69:	gb_est39:*		
26:	gb_est7:*	70:	gb_est40:*		
27:	gb_est8:*	71:	gb_est41:*		
28:	gb_est9:*	72:	gb_est42:*		
29:	gb_est10:*	73:	gb_est43:*		
30:	gb_est11:*	74:	gb_est44:*		
31:	gb_est12:*	75:	em_est31:*		
32:	gb_est13:*	76:	em_est32:*		
33:	gb_est14:*	77:	em_est33:*		
34:	gb_est15:*	78:	em_est34:*		
35:	gb_est16:*	79:	gb_est45:*		
36:	gb_est17:*	80:	gb_est46:*		
37:	gb_est18:*	81:	gb_est47:*		
38:	gb_est19:*	82:	gb_gss1:*		
39:	gb_est20:*	83:	gb_gss2:*		
40:	gb_est21:*	84:	gb_gss3:*		
41:	gb_est22:*	85:	gb_gss4:*		
42:	gb_est23:*	86:	em_gss1:*		
43:	gb_est24:*	87:	em_gss2:*		
44:	gb_est25:*	88:	em_gss3:*		
		89:	em_gss4:*		
		90:	gb_gss5:*		
		91:	gb_gss6:*		
		92:	gb_gss7:*		
		93:	gb_gss8:*		

Result   Query

8   SUMMARIES
           Query

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 1

No.		Score	Match	Length	DB	ID	Description
1		66	100.0	435	23	H38363	H38363 yp50g04.r1
2		30	45.5	306	35	AA586761	AA586761 nn71c10.s
3	c	27.2	41.2	487	22	R62432	R62432 yg52e11.s1
4	c	26.4	40.0	252	25	N87424	N87424 L3865F Huma
5	c	26.2	39.7	835	69	AU080921	AU080921 AU080921
6		26	39.4	254	59	AV080438	AV080438 AV080438
7	c	25.8	39.1	341	28	AA125015	AA125015 mg73c06.r
8	c	25.8	39.1	341	30	AA212205	AA212205 mu80b06.r
9	c	25.8	39.1	394	105	AQ338519	AQ338519 HS_3118_B
10	c	25.8	39.1	416	26	W30350	W30350 mc25e10.r1
11	c	25.8	39.1	426	37	AA726779	AA726779 vu42f03.r
12	c	25.8	39.1	434	28	AA103674	AA103674 mo40c10.r
13		25.8	39.1	438	34	AA466804	AA466804 vd90g11.r
14	c	25.8	39.1	497	27	AA000659	AA000659 mg26a04.r
15	c	25.8	39.1	505	38	AA762271	AA762271 vw60g01.r
16		25.8	39.1	551	38	AA123801	AA123801 mp96d07.r
17	c	25.2	38.2	462	107	AQ433555	AQ433555 HS_5060_A
18		25	37.9	567	108	AQ572153	AQ572153 HS_2108_B
19		24.8	37.6	361	41	AU016229	AU016229 AU016229
20	c	24.8	37.6	382	79	AW261956	AW261956 xq29d10.x
21		24.8	37.6	424	85	AQ718452	AQ718452 HS_5516_B
22	c	24.8	37.6	479	79	AW261950	AW261950 xq29c10.x
23		24.8	37.6	509	49	AI655503	AI655503 tt13e07.x
24	c	24.8	37.6	534	49	AI655522	AI655522 tt13g07.x
25		24.8	37.6	537	49	AI652683	AI652683 wb30e10.x
26	c	24.8	37.6	551	108	AQ541860	AQ541860 RPCI-11-3
27	c	24.6	37.3	235	73	AV345313	AV345313 AV345313
28		24.6	37.3	398	109	AQ522265	AQ522265 HS_5200_A
29		24.4	37.0	663	109	AQ656077	AQ656077 Sheared D
30	c	24.4	37.0	467	23	H17411	H17411 ym40c11.s1
31	c	24.2	36.7	191	49	AI645503	AI645503 vt91h1.x
32	c	24.2	36.7	253	37	AA720125	AA720125 33318 Lam
33	c	24.2	36.7	461	91	AQ861085	AQ861085 nbeb00016E
34		24.2	36.7	497	43	AI195110	AI195110 u160g06.x
35		24.2	36.7	497	23	H55902	H55902 yr02c08.s1
36	c	24	36.4	501	27	AA038832	AA038832 m195g12.r
37		24	36.4	628	41	AI003549	AI003549 a188a09.s
38		23.8	36.1	269	32	AA376236	AA376236 EST88631
39		23.8	36.1	276	69	AV214670	AV214670 AV214670
40		23.8	36.1	440	24	H69538	H69538 yr89c05.s1
41		23.8	36.1	461	25	N80340	N80340 yz91h06.r1
42		23.8	36.1	535	109	AQ655614	AQ655614 Sheared D
43		23.8	36.1	536	85	AQ708302	AQ708302 HS_5569_A
44		23.8	36.1	565	24	N38130	N38130 19357 Lambd
45		23.8	36.1	733	62	AU077466	AU077466 AU077466

ALIGNMENTS

RESULT 1

```
LOCUS       H38363        435 bp   mRNA           EST       16-AUG-1995
DEFINITION  yp50g04.r1 Soares retina N2b4HR Homo sapiens cDNA clone
            IMAGE:190902 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
            RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   H38363
VERSION     H38363.1  GI:907862
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 435)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
```

```
COMMENT     On May 5, 1995 this sequence version replaced gi:797807.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2663
            High quality sequence stops: 276
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2663   Std Error: 0.00
            Seq primer: M13rev
            High quality sequence stop: 276.
FEATURES             Location/Qualifiers
     source          1..435
                     /organism="Homo sapiens"
                     /db_xref="GDB:384731"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:190902"
                     /clone_lib="Soares retina N2b4HR"
                     /sex="male"
                     /tissue_type="retina"
                     /dev_stage="55 year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGAGCGACGCGGCCGCCCTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). The retinas were obtained from a 55 year old
                     Caucasian and total cellular poly(A)+ RNA was extracted 6
                     hrs after their removal. The retina RNA was kindly
                     provided by Roderick R. McInnes M.D. Ph.D. from the
                     University of Toronto. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      109 a    108 c    103 g    109 t    6 others
ORIGIN
```

```
Qy   61 catggg 66
         ||||||
Db  214 CATGGG 219
```

```
Query Match              100.0%;   Score 66;   DB 23;   Length 435;
Best Local Similarity    100.0%;   Pred. No. 4.1e-14;
Matches   66  Conservative   0  Mismatches   0  Indels   0  Gaps   0;
```

```
Qy    1 gtcaatctactggattcaaaaaacaaattcaaaggggagctgggctggattctcttatccatca 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  154 GTCAATCTACTGGATTCAAAAAACAAATTCAAAGGGGAGCTGGGCTGGATTCTCTTATCCATCA 213
```

RESULT 2

```
LOCUS       AA586761       306 bp   mRNA           EST       12-SEP-1997
DEFINITION  nn71c10.s1 NCI_CGAP_Lar1 Homo sapiens cDNA clone IMAGE:1089330 3',
            mRNA sequence.
ACCESSION   AA586761
VERSION     AA586761.1  GI:2397575
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 306)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
```

RESULT 2

```
LOCUS       AA586761/c
DEFINITION  nn71c10.s1 NCI_CGAP_Lar1 Homo sapiens cDNA clone IMAGE:1089330 3',
            mRNA sequence.
```

COMMENT
On Nov 29, 1993 this sequence version replaced gi:430058.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 298.

FEATURES   Location/Qualifiers
source   1..306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1089330"
/clone_lib="NCI_CGAP_Lar1"
/tissue_type="larynx"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: larynx; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
T3. Average insert size: 0.9 kb."

BASE COUNT   93 a   77 c   62 g   74 t
ORIGIN

```
Qy   13 gattcaaaaacaattcaaaggggagctgggggatctcttatccatcacatggg 66
         |||||||| |||||| |||||||||||||| |||||||| ||||| |||||||
Db  237 GATTCAAGAAACACATCTCTGGGGTCTGGGACTGGGACTCTTTTCCATAAACATGAG 184
```

Query Match   Score 30; DB 35; Length 306;
Best Local Similarity   72.2%;   Pred. No. 0.55;
Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

RESULT 3
LOCUS   R62432   487 bp   mRNA   EST   26-MAY-1995
DEFINITION   yg52e11.s1 Soares infant brain 1NIB Homo sapiens cDNA clone
IMAGE:36023 3', mRNA sequence.
ACCESSION   R62432
VERSION   R62432.1   GI:834311
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 487)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   On May 9, 1995 this sequence version replaced gi:804377.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 2443
High quality sequence stops: 337 Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL ; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 2443 Std Error: 0.00

Seq primer: Promega -21m13
High quality sequence stop: 337.

FEATURES   Location/Qualifiers
source   1..487
/organism="Homo sapiens"
/db_xref="GDB:408524"
/db_xref="taxon:9606"
/clone="IMAGE:36023"
/clone_lib="Soares infant brain 1NIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5'
AACTGGAAGAATTCGGCACGAGGAATTTTTTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT   144 a   88 c   108 g   140 t   7 others
ORIGIN

```
Qy    4 aatctactggattcaaaaacaattcaaagggggagctggtctggagatctcttatccatc 59
         ||||| |||| |||||| ||||||||||||||||| ||| |||||||||||||||||| |
Db  332 AATCCACAGTTTCAATTAAATTACAGAGTGGAACAAGGCTGGGTGTGGTGTTCTCCAAC 387
```

Query Match   Score 27.2; DB 22; Length 487;
Best Local Similarity   67.9%;   Pred. No. 5.8;
Matches 38; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

RESULT 4
N87424/c
LOCUS   N87424   252 bp   mRNA   EST   01-APR-1996
DEFINITION   L3865F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
clone L3865 5' similar to RECEPTOR PROTEIN-TYROSINE KINASE (HEK11),
mRNA sequence.
ACCESSION   N87424
VERSION   N87424.1   GI:1440626
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 252)
AUTHORS   Liew,C.C.
TITLE   cDNAs from fetal heart (1996)
JOURNAL   Unpublished (1996)
COMMENT   On May 9, 1995 this sequence version replaced gi:802248.
Contact: Liew CC
Department of Laboratory Medicine and Pathobiology
University of Toronto
Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 4169788758
Fax: 4169785650
Email: liewcc@utcc.utoronto.ca

FEATURES   Location/Qualifiers
source   1..252
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="L3865"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."
Seq primer: GAAATTAACCCTCACTAAAGGG.

RESULT 5

```
BASE COUNT      64 a      56 c      58 g      74 t
ORIGIN
```

```
Qy          7 ctactggattcaaaaacaattcaagggggagctgggcgtgggatctcttatccatcacatggg 66
              || ||||||||||||||||||||||||  |||  ||   |||   |||  ||  ||  ||
Db         96 CTGCTGGATTCTAAAGCACACAACAACAAACAGAGTTGGAGTGGGATTTCCTCCCACCAATGGG 37
```

Query Match   40.0%;   Score 26.4; DB 25; Length 252;
Best Local Similarity 65.0%;   Pred. No. 11;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
LOCUS       AU080921       835 bp    mRNA            EST       20-OCT-1999
DEFINITION  AU080921 Sugano mouse brain mncb Mus musculus cDNA clone MNCb-6241
            5', mRNA sequence.
ACCESSION   AU080921
VERSION     AU080921.1  GI:6085675
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 835).
AUTHORS     Hashimoto,K., Kusuda,J., Toyoda,A., Tanuma,R., Ito,A., Hirata,M.,
            Suzuki,Y., Sasaki,M. and Sugano,S.
TITLE       Isolation of full-length cDNA clones from a mouse brain cDNA
            library made by oligo-capping method
JOURNAL     Unpublished (1999)
COMMENT     On May 18, 1998 this sequence version replaced gi:3137227.
            Contact: Katsuyuki Hashimoto
            National Institute of Infectious Diseases
            Division of Genetic Resources
            23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            Email: khashi@nih.go.jp
            URL: http://www.nih.go.jp/yoken/genbank/.
FEATURES             Location/Qualifiers
     source          1..835
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="MNCb-6241"
                     /clone_lib="Sugano mouse brain mncb"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="TOP10"
                     /note"Organ: brain; Vector: pME18S-FL3; 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCGTTTTTTTTTTTTTTTTTT] double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3.
                     XhoI sites just outside the DraIII sites can be used to
                     isolate the cDNA insert. Size selection was performed to
                     exclude fragments <1.5 kb. Library was constructed by
                     Sugano et al. (University of Tokyo, Institute of Medical
                     Science). Custom primer for sequencing: 5' end primer
                     [CTTCTGCCTCTAAAAGCTGCG]"
BASE COUNT    202 a    230 c    190 g    31 others
ORIGIN
```

RESULT 6

```
Qy          1 gtcaatcttactggattcaaaaacaattcaagggggagctggggggctggat 47
              || || ||| |||||||||||||| |||| ||||||||||  ||||| |||||
Db        147 GTTACTTTATTGGATTCCAGATCTGTTCAGGAGGAGAGCTTGGGTGGAT 193
```

Query Match   39.7%;   Score 26.2; DB 69; Length 835;
Best Local Similarity 72.3%;   Pred. No. 14;
Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
LOCUS       AV080438/c       254 bp    mRNA            EST       25-JUN-1999
DEFINITION  AV080438 Mus musculus stomach C57BL/6J adult Mus musculus cDNA
            clone 2210414F24, mRNA sequence.
ACCESSION   AV080438
VERSION     AV080438.1  GI:5211886
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 254)
AUTHORS     Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
            Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
            Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
            Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
            Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shirai,T., Sogabe,Y.,
            Sugahara,Y., Suzuki,H., Suzuki,Y., Yagame,M., Yamamura,T., Yokota,T.,
            Tominaga,N., Watanabe,S., Yoshino,M., Okazaki,Y. and Hayashizaki,Y.
            Yoshino,M., Muramatsu,M.,
TITLE       RIKEN Mouse ESTs
JOURNAL     Unpublished (1999)
COMMENT     On Mar 10, 1998 this sequence version replaced gi:2948178.
            Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-res@rtc.riken.go.jp
            This clone is available from RIKEN. For request, please use URL
            (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1..254
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2210414F24"
                     /clone_lib="Mus musculus stomach C57BL/6J adult"
                     /sex="male"
                     /tissue_type="stomach"
                     /dev_stage="adult"
BASE COUNT     63 a     62 c     58 g     70 t     1 others
ORIGIN
```

RESULT 7

```
Qy          8 tactggattcaaaaacaattcaagggggagctgggcgtgggatctcttatccatcacatggg 66
              || ||| ||| |||||||| |  ||||||| ||||||| ||| ||| | | ||| || ||||
Db         85 TGCAGGAATCAAAAATCTCTTGTATAGGGCGTGGGTTGCAGCCCTANTCTCACACAGAGG 27
```

Query Match   39.4%;   Score 26; DB 59; Length 254;
Best Local Similarity 64.4%;   Pred. No. 16;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
LOCUS       AA125015/c       341 bp    mRNA            EST       09-FEB-1997
DEFINITION  mg73c06.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
            clone IMAGE:584362 5', mRNA sequence.
ACCESSION   AA125015
VERSION     AA125015.1  GI:1684192
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 341)
```

AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
          Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
          Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
          Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
          Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Sep 12, 1996 this sequence version replaced gi:1297528.
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: mouseest@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:359010
          Seq primer: -28m13 rev1 ET from Amersham
          High quality sequence stop: 294.
FEATURES   Location/Qualifiers
     source   1..341
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:584362"
                      /clone_lib="Stratagene mouse melanoma (#937312)"
                      /tissue_type="melanoma"
                      /dev_stage="M2 cells"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: skin; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
                      Oligo dT. From M2 cells, a highly metastatic derivative of
                      the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
                      Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG
                      3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT   73 a   95 c   88 g   85 t
ORIGIN

Query Match   39.1%; Score 25.8; DB 28; Length 341;
Best Local Similarity   63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   6 tctactggattcaaaaacaatcaaggggagctggcgtcggatctcttatccatcacatgg 65
          |||||| ||| |||| |||||||||||||||||| |||||||| || |||||||| |
Db   75 TCTACGTGATGCAACTACAGAGACAAAAGGGGTGGGGCTTGGCCACAGCACTATCACAATG 16

Qy   66 g 66
          |
Db   15 G 15

RESULT 8
AA212205
LOCUS   AA212205   364 bp   mRNA   EST   31-JAN-1997
DEFINITION mu80b06.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
            clone IMAGE:651827 5' similar to TR:G1184951 G1184951
            PHOSPHOTYROSINE INDEPENDENT LIGAND P62B FOR THE LCK SH2 DOMAIN
            B-CELL ISOFORM ;, mRNA sequence.
ACCESSION   AA212205
VERSION   AA212205.1 GI:1810850
KEYWORDS   EST.
SOURCE   Mus musculus
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 364)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project

JOURNAL   Unpublished (1996)
COMMENT   On May 18, 1995 this sequence version replaced gi:811160.
          Contact: Marra M/Mouse EST Project
          WashU-HHMI Mouse EST Project
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: mouseest@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:397675
          Seq primer: -28m13 rev1 ET from Amersham
          High quality sequence stop: 287.
FEATURES   Location/Qualifiers
     source   1..364
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:651827"
                      /clone_lib="Stratagene mouse melanoma (#937312)"
                      /tissue_type="melanoma"
                      /dev_stage="M2 cells"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: skin; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
                      Oligo dT. From M2 cells, a highly metastatic derivative of
                      the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
                      Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG
                      3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTT 3'"
BASE COUNT   80 a   92 c   101 g   91 t
ORIGIN

Query Match   39.1%; Score 25.8; DB 30; Length 364;
Best Local Similarity   63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   6 tctactggattcaaaaacaatcaaggggagctggcgtcggatctcttatccatcacatgg 65
          |||||| ||| |||| |||||||||||||||||| |||||||| || |||||||| |
Db   197 TCTACGTGATGCAACTACAGAGACAAAAGGGGTGGGGCTTGGCCACAGCACTATCACAATG 138

Qy   66 g 66
          |
Db   137 G 137

RESULT 9
AQ338519
LOCUS   AQ338519   394 bp   DNA   GSS   12-JAN-1999
DEFINITION HS_3118_B1_C10_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3118 Col=19 Row=F, genomic survey
            sequence.
ACCESSION   AQ338519
VERSION   AQ338519.1 GI:4144535
KEYWORDS   GSS.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 394)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE   Sequence-tagged connectors: A sequence approach to mapping and
            sequencing the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887

FEATURES
    source

        Email: jwallace@u.washington.edu
        Clones may be purchased from Research Genetics (info@resgen.com).
        BAC end Web Server: http://www.htsc.washington.edu
        Plate: 3118 row: F column: 19
        Seq primer: M13 Reverse
        Class: BAC ends
        High quality sequence stop: 394.
            Location/Qualifiers
            1..394
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="Plate=3118 Col=19 Row=F"
            /clone_lib="CIT Approved Human Genomic Sperm Library D"
            /sex="male"
            /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
            E-Coli DH10B"

BASE COUNT   87 a   115 c   97 g   95 t
ORIGIN

Qy   65   g 65

Db   158   G 158

Qy   65   g 65

Db   98   ATGTGTGGTTGAACAACACAAACCACACGTGGTGCTGCGTGGACTCCCCGCCTGTGCCCTG 157

Db   98 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||

Qy   5   atctactggattcaaaaacaattcaagggagctggcgtggacctcttatccatcacatg 64

Db   105 |||||||||||||||||||||| ||||| | |||||||||||||||||||||||| ||

Qy   5   Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:349578 5' similar to PIR:S33561 S33561 ret(2)P protein -
fruit fly ;, mRNA sequence.

ACCESSION   W30350
VERSION   W30350.1   GI:1310500
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1   (bases 1 to 416)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
              Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
              Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
              Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
              Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Apr 14, 1993 this sequence version replaced gi:692855.
              Contact: Marra M/Mouse EST Project
              WashU-HHMI Mouse EST Project
              Washington University School of Medicinep
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:221378
              Putative full length read
              Seq primer: ETprimer
              High quality sequence stop: 414.
FEATURES
    source
            Location/Qualifiers
            1..416
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:349578"

RESULT 10
LOCUS   W30350
DEFINITION   mc25e10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
              IMAGE:349578 5' similar to PIR:S33561 S33561 ret(2)P protein -

Query Match   39.1%;   Score 25.8;   DB 105;   Length 394;
Best Local Similarity   63.9%;   Pred. No. 19;   Matches 22;   Indels 0;   Gaps 0;
Matches 39;   Conservative 0;   Mismatches 22;   Indels 0;   Gaps 0;

Qy   5   atctactggattcaaaaacaattcaagggagctggcgtggacctcttatccatcacatg 64

FEATURES
    source

            /clone_lib="Soares mouse p3NMF19.5"
            /dev_stage="19.5 dpc total fetus"
            /lab_host="DH10B (ampicillin resistant)"
            /note="Vector: pT7T3D (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5'
            TGTACCAATCTGAAGTGGGAGCGCGGCCGCCATTTTTTTTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI
            adapters (Pharmacia), digested with Not I and cloned into
            the Not I and Eco RI sites of a modified pT7T3 vector
            (Pharmacia). Library went through one round of
            normalization to a Cot = 5. Library constructed by Bento
            Soares and M.Fatima Bonaldo. RNA was kindly provided by
            Dr. Minoru Ko (Wayne State University)."

BASE COUNT   86 a   111 c   115 g   104 t
ORIGIN

Qy   66   g 66

Db   153   G 153

Qy   66   g 66

Db   213   TCTACGTGATGCAACTACACAGACACAAAAGGGGGGTGGGGGCTTGGCCACAGCACTATCACAATG 154

Qy   6   tctactggattcaaaaacaattcaagggagctggcgtggacctcttatccatcacatgg 65

Db   213 ||||||||||||||||||||| || |||| | |||||||||||||||||||| ||||||

Query Match   39.1%;   Score 25.8;   DB 26;   Length 416;
Best Local Similarity   63.9%;   Pred. No. 19;   Matches 22;   Indels 0;   Gaps 0;
Matches 39;   Conservative 0;   Mismatches 22;   Indels 0;   Gaps 0;

LOCUS   AA726779   426 bp   mRNA   EST   02-JAN-1998
DEFINITION   vu42f03.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
              IMAGE:1194077 5' similar to TR:Q13502 Q13502 PHOSPHOTYROSINE
              INDEPENDENT LIGAND P62B FOR THE LCK SH2 DOMAIN B-CELL ISOFORM ;,
              mRNA sequence.

ACCESSION   AA726779
VERSION   AA726779.1   GI:2744486
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1   (bases 1 to 426)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
              Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
              Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
              Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
              Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On May 5, 1995 this sequence version replaced gi:797862.
              Contact: Marra M/Mouse EST Project
              WashU-HHMI Mouse EST Project
              Washington University School of Medicinep
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:641173
              Seq primer: -28m13 rev2 ET from Amersham
              High quality sequence stop: 419.
FEATURES
    source
            Location/Qualifiers
            1..426
            /organism="Mus musculus"
            /strain="C3H"
            /db_xref="taxon:10090"
            /clone="IMAGE:1194077"
            /clone_lib="Barstead mouse myotubes MPLRB5"

RESULT 11
LOCUS   AA726779/c

RESULT 12

LOCUS   AA103674   434 bp   mRNA   EST   29-OCT-1996
DEFINITION mo40c10.r1 Life Tech mouse embryo 15 5dpc 10667012 Mus musculus
            cDNA clone IMAGE:556050 5' similar to TR:G1184951 G1184951
            PHOSPHOTYROSINE INDEPENDENT LIGAND P62B FOR THE LCK SH2 DOMAIN
            B-CELL ISOFORM ;, mRNA sequence.
ACCESSION   AA103674
VERSION   AA103674.1 GI:1649894
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 434)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Nov 29, 1993 this sequence version replaced gi:430261.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicinep
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:336842
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 416.
FEATURES   Location/Qualifiers
     source   1..434
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:556050"
                     /clone_lib="Life Tech mouse embryo 15 5dpc 10667012"

BASE COUNT   99 a   115 c   109 g   103 t
ORIGIN

Query Match   39.1%; Score 25.8; DB 37; Length 426;
Best Local Similarity   63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   6 tctactggattcaaaaacaattcaagggaggagctgggctggatctctcttattccatcacatgg 65
          ||||| ||||||| |||| ||||| ||| |||| ||||||||||||| |||||||||||||||||
Db   202 TCTACGTGATGCAACTACAAGACACAAAAGGGGGTGGGGCTTGTCCACAGCACTATCACAATG 143

Qy   66 g 66
          | |
Db   142 G 142

RESULT 13

LOCUS   AA466804   438 bp   mRNA   EST   11-JUN-1997
DEFINITION vd90g11.r1 Soares mouse NbMH Mus musculus cDNA clone IMAGE:807908
            5' similar to TR:G1184951 G1184951 PHOSPHOTYROSINE INDEPENDENT
            LIGAND P62B FOR THE LCK SH2 DOMAIN B-CELL ISOFORM ;, mRNA sequence.
ACCESSION   AA466804
VERSION   AA466804.1 GI:2192944
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 438)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Jan 25, 1995 this sequence version replaced gi:637879.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicinep
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:484252
            Putative full length read
            vector to vector length is 511
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 409.
FEATURES   Location/Qualifiers
     source   1..438
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:807908"
                     /clone_lib="Soares mouse NbMH"
                     /sex="male"
                     /tissue_type="heart"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Eco RI; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGACGGCGGCGGGCCGCCGGAAAGTTTTTTTTTTTTTTT
                     TTTTTTTTTTTTTTTTTTTT

BASE COUNT   101 a   108 c   120 g   105 t
ORIGIN

Query Match   39.1%; Score 25.8; DB 28; Length 434;
Best Local Similarity   63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   6 tctactggattcaaaaacaattcaagggaggagctgggctggatctctcttattccatcacatgg 65
          ||||| ||||||| |||| ||||| ||| |||| ||||||||||||| |||||||||||||||||
Db   268 TCTACGTGATGCAACTACAAGACACAAAAGGGGGTGGGGCTTGTCCACAGCACTATCACAATG 209

Qy   66 g 66
          | |
Db   208 G 208

RESULT 14

LOCUS   AA000659   497 bp   mRNA   EST   18-JUL-1996
DEFINITION mg26aa04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:424878 5' similar to PIR:S33561 S33561 ref(2)P protein
            - fruit fly ;, mRNA sequence.
ACCESSION   AA000659
VERSION   AA000659.1   GI:1436522
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 497)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   On Apr 14, 1993 this sequence version replaced gi:693679.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:259430
FEATURES   Location/Qualifiers
     source   1..497
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:424878"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I - Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGCGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'), on equal amounts of mRNA from 2 13.5dpc and 2

BASE COUNT   95 a   127 c   113 g   103 t
ORIGIN

Qy   6 tctactggattcaaaaacaattcaaggggagctgggggatcctcttatccatcaatgg 65
              |||| |||||||||| |||| |||| ||||||||||||||||||||||| |||||||||
Db   160 TCTTACGTGATGCAACTACAAGACAAAAGGGGTGGGGGCTTGGCCACAGCACTATCACAATG 101

Qy   66 g 66
              |
Db   100 G 100

Query Match   39.1%; Score 25.8; DB 34; Length 438;
Best Local Similarity 63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels 0; Gaps 0;

3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the modified pT7T3 vector. RNA
provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

RESULT 15

LOCUS   AA762271   505 bp   mRNA   EST   27-JAN-1998
DEFINITION vw60g01.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1248240 5' similar to TR:O13502 O13502 PHOSPHOTYROSINE
            INDEPENDENT LIGAND P62B FOR THE LCK SH2 DOMAIN B-CELL ISOFORM ;,
            mRNA sequence.
ACCESSION   AA762271
VERSION   AA762271.1   GI:2812018
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 505)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   On May 8, 1995 this sequence version replaced gi:801194.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:661928
FEATURES   Location/Qualifiers
     source   1..505
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1248240"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not

BASE COUNT   129 a   133 c   140 g   94 t   1 others
ORIGIN

Qy   6 tctactggattcaaaaacaattcaaggggagctgggggatcctcttatccatcaatgg 65
              |||| |||||||||| |||| |||| |||||||||||||||||||||||||| |||||||
Db   447 TCTACGTGATGCAACTACAAGACAAAAGGGGTGGGGGCTTGGCCACAGCACTATCACAATG 388

Qy   66 g 66
              |
Db   387 G 387

Query Match   39.1%; Score 25.8; DB 27; Length 497;
Best Local Similarity 63.9%; Pred. No. 19;
Matches 39; Conservative   0; Mismatches 22; Indels 0; Gaps 0;

Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 491.

14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2); double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

```
ORIGIN

            I and cloned into the Not I and Eco RI sites of the
            modified pT7T3 vector.  Library is normalized.  Library
            was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   120 a    132 c    122 g    131 t


Query Match     39.1%;    Score 25.8;   DB 38;   Length 505;
Best Local Similarity  63.9%;   Pred. No. 19;
Matches   39;  Conservative   0;  Mismatches   22;  Indels   0;  Gaps   0;

Qy       6  tctactggattcaaaaacaattcaagggggagctggctggatctcttattccatcacatgg  65
            ||||| ||| |||| |  |||| |  |  |  |  ||| || |  | || | |||||||||
Db     132  TCTACGTGATGCAACTACAAGACACAAAAGGGGGTGGGGGCTTGGCCACAGCACTATCACAATG  73

Qy      66  g  66
            |  |
Db      72  G  72


Search completed: May 15, 2000, 11:20:19
Job time: 19142 sec
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

```
                              GenCore version 4.5
                         Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: May 15, 2000, 12:05:03 ; Search time 102.62 Seconds
                                          (without alignments)
                                          81.662 Million cell updates/sec

Title: US-09-104-340-7
Perfect score: 66
Sequence: 1 gtcaactactggattcaaa..............tctcttatccatcacatggg 66
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 226296 seqs, 63486255 residues

Total number of hits satisfying chosen parameters: 452592

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1:/cgn2_6/ptodata/2/1na/5A_COMB.seq:*
                 2:/cgn2_6/ptodata/2/1na/5B_COMB.seq:*
                 3:/cgn2_6/ptodata/2/1na/5C_COMB.seq:*
                 4:/cgn2_6/ptodata/2/1na/5D_COMB.seq:*
                 5:/cgn2_6/ptodata/2/1na/6_COMB.seq:*
                 6:/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:*
                 7:/cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	100.0	100.0	3132	1	US-08-167-919A-9	Sequence 9, Appli
2	66	66	3132	5	US-08-715-106-9	Sequence 9, Appli
3	59.6	90.3	3254	1	US-08-162-809-15	Sequence 15, Appli
4	34	51.5	3162	4	US-08-449-645A-12	Sequence 12, Appli
5	34	51.5	3162	4	US-08-702-367A-12	Sequence 12, Appli
6	34	51.5	3162	6	PCT-US95-04681-12	Sequence 12, Appli
7	31.2	47.3	4529	4	US-08-449-645A-16	Sequence 16, Appli
8	31.2	47.3	4529	4	US-08-702-367A-16	Sequence 16, Appli
9	31.2	47.3	4529	6	PCT-US95-04681-16	Sequence 16, Appli
10	29.2	44.2	3906	3	US-08-469-537A-102	Sequence 102, Appl
11	29.2	44.2	4165	2	US-08-440-248-1	Sequence 1, Appli
12	29.2	44.2	4165	2	US-08-440-815-1	Sequence 1, Appli
13	27	40.9	2323	6	PCT-US96-00419-6	Sequence 6, Appli
14	27	40.9	2901	6	PCT-US96-00419-4	Sequence 4, Appli
15	27	40.9	2901	6	PCT-US96-00419-1	Sequence 1, Appli
16	26	39.4	4304	6	PCT-US96-00419-1	Sequence 1, Appli
17	25.8	39.1	4322	5	PCT-US96-00419-1	Sequence 1, Appli
18	23.2	35.2	2005	1	US-08-208-108-1	Sequence 1, Appli
19	23.2	35.2	2962	4	US-08-449-645A-10	Sequence 10, Appli
20	23.2	35.2	2962	4	US-08-702-367A-10	Sequence 10, Appli
21	23	34.8	2962	6	PCT-US95-04681-10	Sequence 10, Appli
22	23	34.8	3116	4	US-08-449-645A-14	Sequence 14, Appli
23	23	34.8	3116	4	US-08-702-367A-14	Sequence 14, Appli
24	23	34.8	3116	6	PCT-US95-04681-14	Sequence 14, Appli
25	23	34.8	3348	1	US-08-222-616-34	Sequence 34, Appli
26	22.2	33.6	3592	3	US-08-469-537A-100	Sequence 100, App
27	22.2	33.6	7286	6	PCT-US95-11684-1	Sequence 1, Appli
28	21.8	33.0	867	2	US-08-368-852-14	Sequence 14, Appl
29	21.8	33.0	867	3	US-08-525-940-14	Sequence 14, Appl
30	21.8	33.0	867	4	US-08-976-838-14	Sequence 14, Appl
31	21.8	33.0	2397	4	US-08-525-940-22	Sequence 22, Appl
32	21.8	33.0	2397	4	US-08-976-838-22	Sequence 22, Appl
33	21.8	33.0	2643	3	US-08-525-940-20	Sequence 20, Appl
34	21.8	33.0	2643	4	US-08-976-838-20	Sequence 20, Appl
35	21.8	33.0	2745	3	US-08-525-940-19	Sequence 19, Appl
36	21.8	33.0	2745	4	US-08-976-838-19	Sequence 19, Appl
37	21.8	33.0	2766	3	US-08-525-940-17	Sequence 17, Appl
38	21.8	33.0	2766	4	US-08-976-838-17	Sequence 17, Appl
39	21.8	33.0	3776	1	US-08-162-809-28	Sequence 28, Appl
40	21.6	32.7	4049	1	US-08-162-809-17	Sequence 17, Appl
41	21.6	32.7	4097	1	US-08-162-809-11	Sequence 11, Appl
42	21.4	32.4	372	1	US-08-442-542-5	Sequence 5, Appli
43	21.4	32.4	372	3	US-08-449-287-3	Sequence 3, Appli
44	21.4	32.4	420	3	US-08-449-287-3	Sequence 3, Appli
45	21.4	32.4	777	4	US-08-860-882A-25	Sequence 25, Appl
```

                          ALIGNMENTS

RESULT 1
US-08-167-919A-9
Sequence 9, Application US/08167919A
Patent No. 5674691

; GENERAL INFORMATION:
;   APPLICANT: Boyd, Andrew W.
;   APPLICANT: Simpson, Richard J.
;   APPLICANT: Wicks, Ian
;   APPLICANT: Ward, Larry D.
;   APPLICANT: Wilkinson, David
;   TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
;   TITLE OF INVENTION: AND USE THEREOF
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Scully, Scott, Murphy & Presser
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/167,919A
;   FILING DATE: 18-APR-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PK6841 (AU)
;   FILING DATE: 21-JUN-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PK9992 (AU)
;   FILING DATE: 12-DEC-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/AU92/00294
;   FILING DATE: 19-JUN-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: DiGiglio, Frank S.
;   REGISTRATION NUMBER: 31,346
;   REFERENCE/DOCKET NUMBER: 9159
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 742-4343
;   TELEFAX: (516) 742-4366
;   TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3132 base pairs
;   TYPE: nucleic acid

;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 100..3048
US-08-167-919A-9

Query Match   100.0%; Score 66; DB 1; Length 3132;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 gtcaatctactggattcaaaaacaacaattcaaagggggagctgggggctggacctcttatctccatca 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 187 GTCAATCTACTGGATTCAAAAACAACAATTCAAAGGGGGAGCTGGGGCTGGGACTCTTATTCCATCA 246

Qy   61 catggg 66
         ||||||
Db 247 CAtGGG 252

RESULT 2
US-08-715-106-9
Sequence 9, Application US/08715106
Patent No. 602006
GENERAL INFORMATION:
    APPLICANT: Boyd, Andrew W.
    APPLICANT: Simpson, Richard J.
    APPLICANT: Wicks, Ian
    APPLICANT: Ward, Larry D.
    APPLICANT: Wilkinson, David
    TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
    TITLE OF INVENTION: AND USE THEREOF
    NUMBER OF SEQUENCES: 14
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Scully, Scott, Murphy & Presser
        STREET: 400 Garden City Plaza
        CITY: Garden City
        STATE: New York
        COUNTRY: U.S.A.
        ZIP: 11530
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/715,106
        FILING DATE: 18-SEP-1996
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/167,919
        FILING DATE: 18-APR-1994
        APPLICATION NUMBER: PK6841 (AU)
        FILING DATE: 21-JUN-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PK9992 (AU)
        FILING DATE: 12-DEC-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PCT/AU92/00294
        FILING DATE: 19-JUN-1992
    ATTORNEY/AGENT INFORMATION:
        NAME: Digiglio, Frank S.
        REGISTRATION NUMBER: 31,346
        REFERENCE/DOCKET NUMBER: 9159
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 742-4366
        TELEFAX: (516) 742-4343
        TELEX: 230 901 SANS UR
    INFORMATION FOR SEQ ID NO: 9:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 3132 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 100..3048
US-08-167-919A-9

Query Match   100.0%; Score 66; DB 1; Length 3132;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 gtcaatctactggattcaaaaacaacaattcaaagggggagctgggggctggacctcttatctccatca 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 187 GTCAATCTACTGGATTCAAAAACAACAATTCAAAGGGGGAGCTGGGGCTGGGACTCTTATTCCATCA 246

Qy   61 catggg 66
         ||||||
Db 247 CAtGGG 252

RESULT 3
US-08-162-809-15
Sequence 15, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
    APPLICANT: Pasquale, Elena B.
    APPLICANT: Sajjadi, Fereydoun G.
    TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
    TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
    NUMBER OF SEQUENCES: 26
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: CAMPBELL AND FLORES
        STREET: 4370 La Jolla Village Drive, Suite 700
        CITY: San Diego
        STATE: California
        COUNTRY: United States of America
        ZIP: 92122
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/162,809
        FILING DATE:
        CLASSIFICATION: 514
    ATTORNEY/AGENT INFORMATION:
        NAME: Campbell, Cathryn A.
        REGISTRATION NUMBER: 31,815
        REFERENCE/DOCKET NUMBER: P-LJ 9503
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (619) 535-9001
        TELEFAX: (619) 535-8949
    INFORMATION FOR SEQ ID NO: 15:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 3254 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 32..2980
US-08-162-809-15

Query Match   90.3%; Score 59.6; DB 1; Length 3254;
Best Local Similarity 93.9%; Pred. No. 4.4e-14;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 gtcaatctactggattcaaaaacaacaattcaaagggggagctgggggctggacctcttatctccatca 60

Db   116 GTTAATCTGCTGGATTCAAAAACAATTCAAGGGAGGAGCTGGGCTGGATCTCCTACCCATCA 175
     |||| |||||||||||||||||||||||||||||||||||||||||||||||||| || |||||||

Qy   61 catggg 66
     ||||||

Db   176 CATGGG 181

RESULT 4
US-08-449-645A-12
Sequence 12, Application US/08495645A
Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2976
US-08-449-645A-12

Query Match   51.5%;   Score 34;   DB 4;   Length 3162;
Best Local Similarity   69.7%;   Pred. No. 0.00031;
Matches   46;   Conservative   0;   Mismatches   20;   Indels   0;   Gaps   0;

Qy   1 gtcaatctactggattcaaaaacaattcaagggggagctgggctggatctcttatccatca 60
     || | ||| | || ||||||||||| ||||| | |||| | |||||||| | |||||| |
Db   109 GTGAATTTATTGGATTCACGCACTGTCATGGGGGGGACCTGGGATGGATTGCTTTTCAAAA 168

Qy   61 catggg 66
     ||||||
Db   169 AATGGG 174

RESULT 5
US-08-702-367A-12
Sequence 12, Application US/08702367A
Patent No. 5591246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2976
US-08-702-367A-12

Query Match   51.5%;   Score 34;   DB 4;   Length 3162;
Best Local Similarity   69.7%;   Pred. No. 0.00031;
Matches   46;   Conservative   0;   Mismatches   20;   Indels   0;   Gaps   0;

Qy   1 gtcaatctactggattcaaaaacaattcaagggggagctgggctggatctcttatccatca 60
     || | ||| | || ||||||||||| ||||| | |||| | |||||||| | |||||| |
Db   109 GTGAATTTATTGGATTCACGCACTGTCATGGGGGGGACCTGGGATGGATTGCTTTTCAAAA 168

Qy   61 catggg 66
     ||||||
Db   169 AATGGG 174

RESULT 6
PCT-US95-04681-12
Sequence 12, Application PCT/US9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3162 base pairs

```
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..2976
PCT-US95-04681-12

Query Match    51.5%;  Score 34;  DB 6;  Length 3162;
Best Local Similarity 69.7%;  Pred. No. 0.00031;
Matches 46;  Conservative 0;  Mismatches 20;  Indels 0;  Gaps 0;

Qy    1 gtcaatctactggattcaaaaacaattcaaggggagctggggctggatctcttatccatca 60
          ||      || |||||||||        ||       ||     |||  | ||   ||||||
Db  109 GTGAATTTATTGGATTCACGCACTGTCATGGGGGACCTGGGATGGAATTGCTTTTCCAAAA 168

Qy   61 catggg 66
          ||||||
Db  169 AATGGG 174
```

```
RESULT 7
US-08-449-645A-16
Sequence 16, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
   APPLICANT: Fox, Gary M.
   TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
   TITLE OF INVENTION: Kinases
   NUMBER OF SEQUENCES: 43
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Amgen Patent Operations/RBW
      STREET: 1840 Dehavilland Drive
      CITY: Thousand Oaks
      STATE: California
      COUNTRY: USA
      ZIP: 91320
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/449,645A
      FILING DATE:
      CLASSIFICATION: 435
   ATTORNEY/AGENT INFORMATION:
      NAME: Winter, Robert B.
      REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 4529 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   FEATURE:
      NAME/KEY: CDS
      LOCATION: 186..3182
US-08-449-645A-16

Query Match    47.3%;  Score 31.2;  DB 4;  Length 4529;
Best Local Similarity 70.0%;  Pred. No. 0.0041;
Matches 42;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

Qy    7 ctactggattcaaaaacaattcaaggggagctgggctggatctcttatccatcacatggg 66
          || |||||||  |||   || |||| |       || | |||  | |   ||  ||||||
Db  288 CTGCTGGATTCTAAAGCACAACAAACAGAGTTGGAGTGGATTNCCTCTCCACCCAATGGG 347
```

```
RESULT 8
US-08-702-367A-16
Sequence 16, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
   APPLICANT: Fox, Gary M.
   TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
   TITLE OF INVENTION: Kinases
   NUMBER OF SEQUENCES: 43
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Amgen Patent Operations/RBW
      STREET: 1840 Dehavilland Drive
      CITY: Thousand Oaks
      STATE: California
      COUNTRY: USA
      ZIP: 91320
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/702,367A
      CLASSIFICATION: 435
      FILING DATE:
   ATTORNEY/AGENT INFORMATION:
      NAME: Winter, Robert B.
      REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 4529 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   FEATURE:
      NAME/KEY: CDS
      LOCATION: 186..3182
US-08-702-367A-16

Query Match    47.3%;  Score 31.2;  DB 4;  Length 4529;
Best Local Similarity 70.0%;  Pred. No. 0.0041;
Matches 42;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

Qy    7 ctactggattcaaaaacaattcaaggggagctgggctggatctcttatccatcacatggg 66
          || |||||||  |||   || |||| |       || | |||  | |   ||  ||||||
Db  288 CTGCTGGATTCTAAAGCACAACAAACAGAGTTGGAGTGGATTNCCTCTCCACCCAATGGG 347
```

```
RESULT 9
PCT-US95-04681-16
Sequence 16, Application PC/TUS9504681
GENERAL INFORMATION:
   APPLICANT: Fox, Gary M.
   TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
   TITLE OF INVENTION: Kinases
   NUMBER OF SEQUENCES: 28
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Amgen Patent Operations/RBW
      STREET: 1840 Dehavilland Drive
      CITY: Thousand Oaks
      STATE: California
      COUNTRY: USA
      ZIP: 91320
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US95/04681
      FILING DATE:
```

```
RESULT 10
US-08-469-537A-102
Sequence 102, Application US/0846953TA
Patent No. 5843749
GENERAL INFORMATION:
 APPLICANT: Maisonpierre, et al.
 TITLE OF INVENTION: EHK AND ROR TYROSINE
 TITLE OF INVENTION: KINASES
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: Regeneron Pharmaceuticals, Inc.
  STREET: 777 Old Saw Mill River Road
  CITY: Tarrytown
  STATE: NY
  COUNTRY: U.S.A.
  ZIP: 10591
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSEQ Version 2.0
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/469,537A
  FILING DATE: 06-JUN-1995
  CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: USSN 08/406,247
  FILING DATE: 17-MAR-1995
  APPLICATION NUMBER: USSN 08/144,992
  FILING DATE: 28-OCT-1993
  APPLICATION NUMBER: USSN 07/736,559
  FILING DATE: 26-JUL-1991
 ATTORNEY/AGENT INFORMATION:
  NAME: Kempler, Ph.D., Gail M
  REGISTRATION NUMBER: 32,143
  REFERENCE/DOCKET NUMBER: REG 070C
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: 914-345-7400
  TELEFAX: 914-345-7721
  TELEX:
 INFORMATION FOR SEQ ID NO: 102:
  SEQUENCE CHARACTERISTICS:
   LENGTH: 3906 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: unknown
   TOPOLOGY: unknown
   MOLECULE TYPE: DNA
```

```
Query Match      47.3%;  Score 31.2;  DB 6;  Length 4529;
Best Local Similarity   70.0%;  Pred. No. 0.0041;
Matches  42;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy      7 ctactggattcaaaaacaattcaaggggagctgggctgggcgtgggtatcctcttatcacaatggg 66
          ||||||||||||  |||| ||  ||  |||  |  |||  ||||| ||||||||||| |||||| ||
Db    288 CTGCTGGATTCTAAAGCAACAACAAACAGAGTTGGAGTGGATTTCCTCTCCACCAATGGG 347
```

```
RESULT 11
US-08-442-248-1
Sequence 1, Application US/08442248
Patent No. 5759063
GENERAL INFORMATION:
 APPLICANT: Caras, Ingrid W.
 APPLICANT: Winslow, John W.
 TITLE OF INVENTION: AL-1 Neurotrophic Factor
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
  STREET: 460 Point San Bruno Blvd
  CITY: South San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94080
 COMPUTER READABLE FORM:
  MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: patin (Genentech)
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/442,248
  FILING DATE: 15-MAY-1995
  CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/330128
  FILING DATE: 27-OCT-1994
 ATTORNEY/AGENT INFORMATION:
  NAME: Torchia, Timothy E.
  REGISTRATION NUMBER: 36,700
  REFERENCE/DOCKET NUMBER: 920C4
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-8674
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
   LENGTH: 4165 bases
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
 US-08-442-248-1
```

```
Query Match      44.2%;  Score 29.2;  DB 2;  Length 4165;
Best Local Similarity   65.2%;  Pred. No. 0.023;
Matches  43;  Conservative   0;  Mismatches  23;  Indels   0;  Gaps   0;

Qy      1 gtcaatctactggattcaaaaacaattcaaggggagctgggcgtgggcgtgggatctcttatcatccatca 60
          ||||||||||||||||||||||||| || ||||  |||||| | ||||| ||| ||||| |||| |||||
Db    727 GTGAAATTTATTGGATTCGGCGCACTGTCCTGGGAGGAGAGACCTTGGAGATGGATTGCTTTTTCCAAAG 786
```

```
Qy     61 catggg 66
          |||||
Db    722 AATGGG 727
```

```
FEATURE:
 NAME/KEY: CDS
 LOCATION: 476..3493
US-08-469-537A-102
```

```
Query Match      44.2%;  Score 29.2;  DB 3;  Length 3906;
Best Local Similarity   65.2%;  Pred. No. 0.023;
Matches  43;  Conservative   0;  Mismatches  23;  Indels   0;  Gaps   0;

Qy      1 gtcaatctactggattcaaaaacaattcaaggggagctgggcgtgggcgtgggatctcttatcatccatca 60
          |||||||||||||||||||||||||  ||||||| |||||| |||||| |||||  ||||| ||| ||||
Db    662 GTGAATTATTGGATTCGGCGCACTGTCCTGGGAGGAGAGACCTTGGAGATTGCTTTTTCCAAAG 721
```

```
Qy     61 catggg 66
          |||||
Db    662 GTGAATTTATTGGATTCGGCGCACTGTCCTGGGAGGACCTTGGATGGAATGCTTTTTCCAAAG 721
```

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
 NAME: Winter, Robert B.
 REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 4529 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
 FEATURE:
  NAME/KEY: CDS
  LOCATION: 186..3182
PCT-US95-04681-16
```

```
Query Match      47.3%;  Score 31.2;  DB 6;  Length 4529;
Best Local Similarity   70.0%;  Pred. No. 0.0041;
Matches  42;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy      7 ctactggattcaaaaacaattcaaggggagctgggctgggcgtgggtatcctcttatcacaatggg 66
          ||||||||||||  |||| ||  ||  |||  |  |||  ||||| ||||||||||| |||||| ||
Db    288 CTGCTGGATTCTAAAGCAACAACAAACAGAGTTGGAGTGGATTTCCTCTCCACCAATGGG 347
```

Db   787 AATGGG 792
           ||||||

RESULT 12
US-08-440-815-1
Patent No. 5798448
Sequence 1, Application US/08440815
GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; APPLICANT: Winslow, John W.
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,815
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330128
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 920C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4165 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-440-815-1

Query Match   44.2%; Score 29.2; DB 2; Length 4165;
Best Local Similarity   65.2%; Pred. No. 0.023;
Matches 43; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy   1 gtcaatctactggatccaaaaacaattcaagggggagctgggctggggatctcttatccaatcacatca 60
          ||| || || ||||| || || || || |||| || |||| ||| ||||
Db   727 GTGAATTATTGGATTCGCGCACTGTCCTGGGAGACCTTGGAGGAATTGGAATTGCTTTTCCAAAG 786

Qy   61 catggg 66
          ||||||
Db   787 AATGGG 792

RESULT 13
PCT-US96-00419-6
Sequence 6, Application PCT/US9600419
GENERAL INFORMATION:
; APPLICANT: Thomas Ciossek, Axel Ullrich, Birgit
; APPLICANT: Millauer
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: MDK1 SIGNAL TRANSDUCTION DISORDERS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00419
; FILING DATE: January 3, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: none
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 208/007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2323
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic
PCT-US96-00419-6

Query Match   40.9%; Score 27; DB 6; Length 2323;
Best Local Similarity   66.1%; Pred. No. 0.14;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy   8 tactggattcaaaaacaattcaagggagctgggctggggatctcttatccaatcacatggg 66
          |||||| || | ||||| || || |||| |||||||||||||||||||
Db   336 TACTGGACTCGAAAGCACAACAAACAGAATTGGAATTGGAATGATTTCCTCTCCACCAGTGGG 394

RESULT 14
PCT-US96-00419-4
Sequence 4, Application PCT/US9600419
GENERAL INFORMATION:
; APPLICANT: Thomas Ciossek, Axel Ullrich, Birgit
; APPLICANT: Millauer
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: MDK1 SIGNAL TRANSDUCTION DISORDERS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00419
; FILING DATE: January 3, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application

```
RESULT 15
PCT-US96-00419-1
Sequence 1, Application PC/TUS9600419
GENERAL INFORMATION:
   APPLICANT: Thomas Ciossek, Axel Ullrich, Birgit
   APPLICANT: Millauer
   TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
   TITLE OF INVENTION: MDK1 SIGNAL TRANSDUCTION DISORDERS
   NUMBER OF SEQUENCES: 10
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Lyon & Lyon
      STREET: 633 West Fifth Street
      CITY: Los Angeles
      STATE: California
      COUNTRY: USA
      ZIP: 90071
   COMPUTER READABLE FORM:
      MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
      COMPUTER: IBM compatible
      OPERATING SYSTEM: IBM P.C. DOS (Version 5.1)
      SOFTWARE: WordPerfect (Version 5.0)
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US96/00419
      FILING DATE: January 3, 1995
   PRIOR APPLICATION DATA:
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
   PRIOR APPLICATION DATA: Including application
   PRIOR APPLICATION DATA: described below:
   APPLICATION NUMBER:
      FILING DATE:
   ATTORNEY/AGENT INFORMATION:
      NAME: Warburg, Richard J.
      REGISTRATION NUMBER: 32,327
      REFERENCE/DOCKET NUMBER: 208/007
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (213) 489-1600
      TELEFAX: (213) 955-0440
      TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 4304
      TYPE: nucleic acid
      STRANDEDNESS: single
```

```
                                                              none


Qy     8 tactggattcaaaaacaattcaaggggagctgggctggatctcttatccatcaacatgg 66
         |||||||| || ||| ||| | ||||| |||  | |||| |  |||||||| || |||||
Db   336 TACTGGACTCGAAAGCACAACAAACAGAATTGGAATGGATTTCCTCTCTCCACCCAGTGGG 394

       Best Local Similarity    40.9%;   Score 27;  DB 6;   Length 4304;
       Matches 39;   Conservative   0;  Mismatches  20;  Indels   0;   Gaps   0;

       Query Match
       PCT-US96-00419-1
;      MOLECULE TYPE: nucleic
;      TOPOLOGY: linear
;

Search completed: May 15, 2000, 12:05:06
Job time: 4425 sec
```

```
PCT-US96-00419-4
;      MOLECULE TYPE: nucleic
;      TOPOLOGY: linear
;      STRANDEDNESS: single
       TYPE: nucleic acid
       LENGTH: 2901
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 4:
      TELEX: 67-3510
      TELEFAX: (213) 955-0440
      TELEPHONE: (213) 489-1600
   TELECOMMUNICATION INFORMATION:
      REFERENCE/DOCKET NUMBER: 208/007
      REGISTRATION NUMBER: 32,327
      NAME: Warburg, Richard J.
   ATTORNEY/AGENT INFORMATION:
      FILING DATE:
      APPLICATION NUMBER:
   PRIOR APPLICATION DATA: described below:
   PRIOR APPLICATION DATA: Including application
   PRIOR APPLICATION DATA:
   PRIOR APPLICATION DATA:
      CLASSIFICATION:
   APPLICATION DATA:
   FILING DATE:
APPLICATION NUMBER:
PRIOR APPLICATION DATA:  described below:
   FILING DATE:
APPLICATION NUMBER:
PRIOR APPLICATION DATA:  described below:
```

```
Qy     8 tactggattcaaaaacaattcaaggggagctgggctggatctcttatccatcaacatgg 66
         |||||||| || ||| ||| | ||||| |||  | |||| |  |||||||| || |||||
Db   336 TACTGGACTCGAAAGCACAACAAACAGAATTGGAATGGATTTCCTCTCTCCACCCAGTGGG 394

       Best Local Similarity    40.9%;   Score 27;  DB 6;   Length 2901;
       Matches 39;   Conservative   0;  Mismatches  20;  Indels   0;   Gaps   0;

       Query Match
```

```
none
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 15, 2000, 12:08:28 ; Search time 111.14 Seconds
(without alignments)
148.575 Million cell updates/sec

Title: US-09-104-340-7
Perfect score: 66
Sequence: 1 gtcaatctactggattcaaa..........tctcttattccatcacatggg 66

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues
Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	66	100.0	3132	1	Q34513	HEK coding sequenc
	2	59.6	90.3	3254	1	Q90659	Eph-related PTK Ce
	3	34	51.5	3162	1	T02947	EPH-like receptor
c	4	31.6	47.9	3663	1	T72320	EPH-like receptor
	5	31.2	47.3	4529	1	T02949	Embryonic stem cel
	6	29.2	44.2	3906	1	V70207	EPH-like receptor
	7	29.2	44.2	4165	1	T18893	Rat receptor tyros
	8	27	40.9	2323	1	V70207	Rat REK7 cDNA. AL-
	9	27	40.9	2901	1	T32961	Mouse developmenta
c	10	27	40.9	2901	1	T32962	Mouse developmenta
	11	26	39.4	4304	1	T32961	Mouse developmenta
c	12	25.8	39.1	4322	1	V58192	Mouse Bsk receptor
	13	23.2	35.2	2005	1	Q71303	Mouse osteoblast-s
	14	23	34.8	2962	1	T02946	EPH-like receptor
	15	23	34.8	3116	1	T02948	EPH-like receptor
c	16	22.4	33.9	3348	1	T03100	Protein tyrosine-k
c	17	22.4	33.9	110000	1	T58840_2	Continuation (3 of
c	18	22.2	33.6	110000	1	T58840_3	Continuation (4 of
	19	22.2	33.6	1000	1	V45121	Mus musculus cerbe
	20	22.2	33.6	3592	1	V70208	Rat receptor tyros
c	21	22.2	33.6	7286	1	T14547	Cytotactin gene. C
	22	22.2	33.6	12923	1	N90338	Sequence of human
	23	22	33.3	1228	1	T04180	scFv-PP coding seq
c	24	22	33.3	1758	1	V84573	Human secreted pro
	25	21.8	33.0	320	1	Q60056	Human brain Expres
	26	21.8	33.0	530	1	N80105	X mutant of L-A ds
	27	21.8	33.0	867	1	T99064	nhrCP coding seque
	28	21.8	33.0	867	1	X01374	T-lymphocyte seque
	29	21.8	33.0	2397	1	X01386	T-lymphocyte prote
	30	21.8	33.0	2643	1	X01385	T-lymphocyte prote
	31	21.8	33.0	2745	1	X01384	T-lymphocyte prote
	32	21.8	33.0	2766	1	X01383	T-lymphocyte prote
	33	21.8	33.0	2819	1	V19910	Human pro-protein
c	34	21.6	32.7	13367	1	X20554	Polynucleotide seq
		21.6	32.7	2187	1	Q12399	Human hepatocyte g
c	35	21.6	32.7	2187	1	Q22146	Complete human HGF
c	36	21.6	32.7	2187	1	V20546	Human recombinant
c	37	21.6	32.7	2187	1	X15186	Nucleic acid encod
c	38	21.6	32.7	2187	1	Q22143	Human HGF gene par
c	39	21.6	32.7	2395	1	Q12397	Human hepatocyte 9
c	40	21.6	32.7	3105	1	T07308	Receptor tyrosine
c	41	21.6	32.7	3105	1	T84528	Mouse Nuk tyrosine
c	42	21.6	32.7	3751	1	Q62461	Human embryonal ki
c	43	21.6	32.7	3776	1	Q90655	Eph-related PTK Ce
c	44	21.6	32.7	4049	1	Q90660	Eph-related PTK Ce
c	45	21.6	32.7	4097	1	Q90657	Eph-related PTK Ce

ALIGNMENTS

RESULT 1

```
ID   Q34513   standard; DNA; 3132 BP.
AC   Q34513;
DT   24-MAY-1993 (first entry)
DE   HEK coding sequence.
KW   HEK coding sequence.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation; ss.
OS   Synthetic.
KW   Synthetic.
FH   Key             Location/Qualifiers
FT   5'utr           1..99
FT                   /*tag= a
FT   cds             100..3051
FT                   /*tag= b
FT   3'utr           3052..3132
FT                   /*tag= c
FT   misc_feature    1723..1795
FT                   /*tag= e
FT                   /note= "potential transmembrane region"
FT   mat_peptide     160..216
FT                   /*tag= f
FT                   /note= "Purified HEK protein #1"
FT   mat_peptide     2617..2679
FT                   /*tag= g
FT                   /note= "Purified HEK protein #2"
FT   signal_peptide  100..159
FT                   /*tag= d
PN   WO9300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU00294.
PR   19-JUN-1992; AU-006841.
PR   21-JUN-1991; AU-006841.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD; Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   P-PSDB; R31466.
DR   WPI; 93-036373/04.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence encodes human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the
CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence 3132 BP; 888 A; 711 C; 759 G; 774 T;
```

Query Match   100.0%; Score 66; DB 1; Length 3132;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 gtcaatctactggattcaaaaaacaattcaaggggggagctgggctggatctctattccatca 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   187 GTCAATCTACTGGATTCAAAAAACAATTCAAGGGGAGCTGGGCTGGATCTCTTATCCATCA 246

Qy    61 catggg 66
           ||||||
Db   247 CATGGG 252
```

RESULT 2
Q90659
ID   Q90659 standard; DNA; 3254 BP.
AC   Q90659;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek4.
KW   Cek4; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis; ss.
OS   Gallus sp.
FH   Key   Location/Qualifiers
FT   cds   32..2980
FT   /*tag= a
PN   WO9515375-A.
PD   08-JUN-1995.
PF   07-SEP-1994; U10140.
PR   03-DEC-1993; US-162809.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PI   Pasquale EB, Sajjadi FG;
DR   WPI; 95-215256/28.
DR   P-PSDB; R75711.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Disclosure; Page 85-89; 129pp; English.
CC   Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and
CC   embryonic tissues, as well as in the adult brain and retina.
SQ   Sequence 3254 BP; 926 A; 737 C; 796 G; 795 T;

Query Match   90.3%; Score 59.6; DB 1; Length 3254;
Best Local Similarity 93.8%; Pred. No. 6.3e-14;
Matches   62; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1 gtcaatctactggattcaaaaaacaattcaaggggggagctgggctggatctctattccatca 60
          | ||||| | |||||||||||||||||||||||||| | |||||||||||||||| ||||||
Db   116 GTTAATCTCGCTGGATTCAAAAAACAATTCAAGGGGAGAGCTGGGCTGGATCTCTACCCATCA 175

Qy    61 catggg 66
          ||||||
Db   176 CATGGG 181
```

RESULT 3
T02947
ID   T02947 standard; cDNA; 3162 BP.
AC   T02947;
DT   16-APR-1996 (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK7 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK7;
KW   human eph-like receptor protein tyrosine kinase; therapy; diagnosis; vector; antibody; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds   1..2976
FT   /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-37799/48.

DR   P-PSDB; R85090.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc. used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 1; Page 49-54; 133pp; English.
CC   New nucleic acid encoding EPH-like receptor protein
CC   tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC   cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC   were isolated from a human foetal brain cDNA library using a directed
CC   PCR approach with primers (see T02960-61) based on conserved regions of
CC   receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC   extensive homology to the catalytic domain of chicken EPH-like
CC   receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC   EPH-like receptor. The isolated cDNAs are used for prodn. of
CC   recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC   to detect abnormalities in HEK receptor genes.
SQ   Sequence 3162 BP; 921 A; 667 C; 775 G; 799 T;

Query Match   51.5%; Score 34; DB 1; Length 3162;
Best Local Similarity 69.7%; Pred. No. 0.00049;
Matches   46; Conservative   0; Mismatches   20; Indels   0; Gaps   0;

```
Qy     1 gtcaatctactggattcaaaaaacaattcaaggggggagctgggctggatctctattccatca 60
          || |||  ||| |||||||||  |  | | || ||||||| |||| ||||| ||| |||||||
Db   109 GTGAATTTATTGGATTCACCGCACTGTCATGGGGGGACCTGGGATGGATTGCTTTTTCCAAAA 168

Qy    61 catggg 66
          | ||||
Db   169 AATGGG 174
```

RESULT 4
T72320
ID   T72320 standard; DNA; 3663 BP.
AC   T72320;
DT   15-SEP-1997 (first entry)
DE   Embryonic stem cell kinase (Esk) DNA.
KW   Embryonic stem cell kinase; receptor tyrosine kinase; mucositis;
KW   epithelium; signal transduction; gene therapy; diagnosis; ss.
OS   Mus musculus.
FH   Key   Location/Qualifiers
FT   cds   355..3288
FT   /*tag= a
FT   signal_peptide 355..432
FT   /*tag= b
FT   mat_peptide   433..3285
FT   /*tag= c
PN   WO9723629-A1.
PD   03-JUL-1997.
PF   20-DEC-1996; AU0826.
PR   22-DEC-1995; AU-007277.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PI   Boyd AW, Lickliter J;
DR   WPI; 97-351065/32.
DR   P-PSDB; W19258.
PT   DNA encoding embryonic stem cell kinase, receptor tyrosine kinase -
PT   for production of modulators and antibodies, useful to treat
PT   mucositis and other disorders involving epithelium
PS   Claim 4; Page 48-54; 98pp; English.
CC   An isolated nucleic acid molecule (T72320) codes for novel mouse
CC   embryonic stem cell kinase (Esk), a receptor tyrosine kinase that
CC   belongs to the Eph subfamily. To isolate the molecule, RNA from
CC   the murine embryonic stem cell line W9.5 was subjected to RT-PCR
CC   using primers (see also T72321-24) based on conserved Eph
CC   sequences. A labelled probe from an isolated clone was used to
CC   screen a mouse liver cDNA library, identifying a single clone that
CC   included the complete coding region of Esk. The gene for murine
CC   Esk is located on bands B1 and B2 of chromosome 6. Esk nucleic
CC   acids can be used to produce recombinant Esk polypeptides, develop
CC   probes, and as sense and antisense molecules for diagnosis or
CC   therapy.
SQ   Sequence 3663 BP; 869 A; 976 C; 1015 G; 801 T;

Query Match
Best Local Similarity 47.9%; Score 31.6; DB 1; Length 3663;
Matches 43; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy    1 gtcaatctactggattcaaaaacaattcaagggagctgggctggatctcttatcatca  60
         |||| ||| |||||||||||||||||||| ||| |||||||||||| || ||| ||
Db  439 GTCACTCTAATGGACACAACAAGCACACACACAAGGAGAGAGCTGGGCGTTCGGATCCCCCA 498

Qy   61 ca  62
         |
Db  499 GA  500
```

RESULT 5

ID T02949 standard; cDNA; 4529 BP.
AC T02949;
DT 16-APR-1996 (first entry)
DE EPH-like receptor protein tyrosine kinase HEK11 cDNA.
KW EPH-like receptor protein tyrosine kinase; PTK; HEK11;
KW human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS Homo sapiens.
FH Key   Location/Qualifiers
FT cds   186..3182
FT   /*tag= a
PN WO9528484-A1.
PD 26-OCT-1995.
PF 14-APR-1995; U04681.
PR 15-APR-1994; US-229509.
PA (AMGE-) AMGEN INC.
PI Fox GM, Jing S, Welcher AA;
PT New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT and related vectors, host cells, proteins, antibodies etc., used
PT diagnostically and therapeutically to modulate receptor activation
PT or prodn.
PS Claim 1; Page 66-71; 133pp; English.
DR P-PSDB; R85092.
DR WPI; 95-373799/48.
CC cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC were isolated from a human foetal brain cDNA library using a directed
CC PCR approach with primers (see T02960-61) based on conserved regions of
CC receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC extensive homology to the catalytic domain of chicken EPH-like
CC receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC EPH-like receptor. The isolated cDNAs are used for prodn. of
CC recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC to detect abnormalities in HEK receptor genes.
SQ Sequence 4529 BP; 1449 A; 834 C; 1000 G; 1246 T;

Query Match
Best Local Similarity 70.0%; Score 31.2; DB 1; Length 4529;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy    7 ctactggattcaaaaacaattcaagggagctgggctggatctcttatcacatggg  66
         ||||||||| |||||||||||||||| || |||||||||||| ||| |||| ||
Db  288 CTGCTGGATTCTAAAGCACAACAAACAGAGAGTTGGAGTGGAGTTCCTTCCACCCAATGGG 347
```

RESULT 6

ID V70207 standard; DNA; 3906 BP.
AC V70207;
DT 11-FEB-1999 (first entry)
DE Rat receptor tyrosine kinase Ehk-1 encoding DNA.
KW Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW binding protein; BDNF; NT-3; diagnosis; ss.
OS Rattus sp.
FH Key   Location/Qualifiers
FT CDS   476..3493
FT   /*tag= a
PN US5843749-A.
PD 01-DEC-1998.
PF 06-JUN-1995; US-469537.
PF 17-MAR-1995; US-406247.
PR 26-JUL-1991; US-736559.
PR 28-OCT-1993; US-144992.
PR 06-JUN-1995; US-469537.
PA (REGE-) REGENERON PHARM INC.
PI Maisonpierre PC, Masiakowski P, Yancopoulos GD;
DR WPI; 99-044584/04.
DR P-PSDB; W83147.
PT DNA encoding receptor tyrosine kinase proteins - and corresponding
PT proteins
PS Claim 5; Fig 22; 194pp; English.
CC The present invention describes nucleic acid molecules for ror-1,
CC ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC tyrosine kinases. The present sequence encodes rat Ehk-1.
SQ Sequence 3906 BP; 1097 A; 876 C; 990 G; 943 T;

```
Qy   61 catggg  66
         |||||
Db  722 AATGGG 727
```

RESULT 7

Query Match
Best Local Similarity 44.2%; Score 29.2; DB 1; Length 3906;
Matches 43; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
Qy    1 gtcaatctactggattcaaaaacaattcaagggagctgggctggatctcttatcatca  60
         | ||| || ||||||||||||||||||||| |||||||||| ||| ||| ||||
Db  662 GTGAATTTATTGGATTCGCGCACTGTCCTGGGAGAGACCTTGGATGGAGATTGCTTTTCCAAAG 721
```

ID T18893 standard; cDNA; 4165 BP.
AC T18893;
DT 05-JAN-1997 (first entry)
DE Rat REK7 cDNA.
KW REK7; eph-related tyrosine kinase receptor; AL-1; ligand;
KW neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis;
KW ss.
OS Rattus sp.
FH Key   Location/Qualifiers
FT cds   541..3327
FT   /*tag= a
FT signal_peptide 541..711
FT   /*tag= b
FT mat_peptide   712..3324
FT   /*tag= c
PN WO9613518-A1.
PD 09-MAY-1996.
PF 26-OCT-1995; U14016.
PR 27-OCT-1994; US-330128.
PR 07-JUN-1995; US-486449.
PA (GETH) GENENTECH INC.
PI Caras IW, Winslow JW;
DR WPI; 96-239448/24.
DR P-PSDB; R97853.
PT AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
PT treatment and diagnosis of neuronal disorders and
PT angiogenesis-related conditions.
PS Example 1; Page 47-49; 75pp; English.
CC A cDNA clone (T18893) codes for rat REK7 (R97853), an eph-related
CC tyrosine kinase receptor, for which AL-1 (see also W9785) is a
CC ligand. It was isolated using degenerate receptor tyrosine kinase
CC primers (T18894-96) to amplify cDNAs of an adult mouse hippocampal
CC cDNA library. A PCR fragment was used as a probe to isolate the
CC full-length REK7 cDNA from a rat hippocampal cDNA library. An
CC REK-IgG fusion was used to screen cultured cell lines for surface

CC   expression of REK7-binding activity. Primers based on isolated
CC   ligands were used to amplify human breast carcinoma BT20 cell
CC   cDNA, and an amplified fragment was used to screen a human foetal
CC   brain cDNA library, leading to the isolation of AL-1 cDNA (T18897).
SQ   Sequence 4165 BP; 1157 A; 907 C; 1035 G; 1066 T;

Query Match   44.2%; Score 29.2; DB 1; Length 4165;
Best Local Similarity   65.2%; Pred. No. 0.038;
Matches 43; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy   1 gtcaatctactggatcaaaaacaattcaagggagctgggctgggatctcttcatca 60
              ||||| ||||| |||||||| ||||| |||| ||||||||||||||| |||||||| |
Db   727 GTGAATTTATTGGATTCGCGCACTGTCCTCGGGGAGACCTTGGGATGGATTGCTTTTCCAAAG 786

Qy   61 catggg 66
              ||||||
Db   787 AATGGG 792

RESULT 8

ID   T32962 standard; cDNA; 2323 BP.
AC   T32962;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1; MDK1 T2 clone.
KW   Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.
FH   Key   Location/Qualifiers
FT   cds   233..2113
FT   /*tag= a
FT   signal_peptide 233..316
FT   /*tag= b
FT   mat_peptide   317..2110
FT   /*tag= c
FT   misc_feature   2017..2113
FT   /*tag= d
FT   /note= "divergent sequence due to alternative
FT   splicing"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; U00419.
PR   03-JAN-1995; US-368776.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
WPI; 96-333988/33.
DR   P-PSDB; W03423.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 115-116; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC   coded for truncated versions (W03422 and W03423, respectively) of
CC   the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC   new member of the eck/eph family of receptor tyrosine kinases.
CC   MDK T1 and T2 each possess the entire ectodomain, the transmembrane
CC   domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC   but lack the catalytic tyrosine kinase domain. They may have a
CC   modulatory function. The cDNA clones can be used to produce MDK1
CC   T1 and T2, which are useful for screening potential agents for
CC   treatment of diseases characterised by abnormal signal transduction.
SQ   Sequence 2323 BP; 680 A; 497 C; 609 G; 537 T;

Qy   8 tactggattcaaaaacaattcaagggaggagctgggctggcatcttcttatccatcacatggg 66
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   8 TACTGGACTCGAAAGCACAACAAACAGAATTGGAAATGGAATTTCCTCTCCACCCAGTGGG 66

RESULT 9

ID   T32961 standard; cDNA; 2901 BP.
AC   T32961;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1; MDK1 T1 clone.
KW   Mouse developmental kinase 1; MDK1 T1; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.
FH   Key   Location/Qualifiers
FT   cds   233..2065
FT   /*tag= a
FT   signal_peptide 233..316
FT   /*tag= b
FT   mat_peptide   317..2062
FT   /*tag= c
FT   misc_feature   2031..2065
FT   /*tag= d
FT   /note= "divergent sequence due to alternative
FT   splicing"
FT   polya_signal   2839..2843
FT   /*tag= e
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; U00419.
PR   03-JAN-1995; US-368776.
PA   (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
WPI; 96-333988/33.
DR   P-PSDB; W03422.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
PS   Example 1; Page 111-12; 128pp; English.
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC   coded for truncated versions (W03422 and W03423, respectively), a
CC   the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC   new member of the eck/eph family of receptor tyrosine kinases.
CC   MDK T1 and T2 each possess the entire ectodomain, the transmembrane
CC   domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC   but lack the catalytic tyrosine kinase domain. They may have a
CC   modulatory function. The cDNA clones can be used to produce MDK1
CC   T1 and T2, which are useful for screening potential agents for
CC   treatment of diseases characterised by abnormal signal transduction.
SQ   Sequence 2901 BP; 890 A; 578 C; 650 G; 783 T;

Qy   8 tactggattcaaaaacaattcaagggaggagctgggctggcatcttcttatccatcacatggg 66
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   8 TACTGGACTCGAAAGCACAACAAACAGAATTGGAAATGGAATTTCCTCTCCACCCAGTGGG 66

RESULT 10

ID   T32960 standard; cDNA; 4304 BP.
AC   T32960;
DT   11-NOV-1996 (first entry)
DE   Mouse developmental kinase 1 cDNA.
KW   Mouse developmental kinase 1; MDK1; receptor tyrosine kinase; RTK;
KW   signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.

Query Match   40.9%; Score 27; DB 1; Length 2901;
Best Local Similarity   66.1%; Pred. No. 0.25;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy   8 tactggattcaaaaacaattcaagggaggagctgggctggcatcttcttatccatcacatggg 66

Db   336 TACTGGACTCGAAAGCACAACAAACAGAATTGGAAATGGAATTTCCTCTCCACCCAGTGGG 394

FH	Key	Location/Qualifiers
FT	cds	233..3229
FT	signal_peptide	233..316
FT		/*tag= a
FT		317..3226
FT		/*tag= b
FT	mat_peptide	317..3226
FT		3230..3628
FT		/*tag= c
FT	3'utr	3230..3628
FT		3568..3573
FT		/*tag= d
FT	polya_signal	3568..3573
FT		3627..4304
FT		/*tag= e
FT	3'utr	3627..4304
FT		/*tag= f
FT		/note= "alternative 3'-untranslated region, found in one isolated clone"

PN WO9621013-A1.
PD 11-JUL-1996.
PF 03-JAN-1995; U00419.
PF 03-JAN-1995; US-368776.
PA (PLAC) MAX PLANCK GES FOERDERING WISSENSCHAFTEN.
PA (SUGE-) SUGEN INC.
PI Clossek T, Millauer B, Ullrich A;
DR WPI; 96-333988/33.
DR P-PSDB: W03421.
PT New mouse development kinase 1 gene - used for developing prods. for
PT diagnosis and treatment of abnormalities in signal transduction
PS pathways
CC Example 1; Page 103-105; 128pp; English.
CC A cDNA clone (T32960) codes for mouse developmental kinase 1 (MDK1)
CC (W03421), a new member of the eck/eph family of receptor tyrosine
CC kinases (RTKs). To isolate the clone, cDNA from mouse embryos was
CC subjected to PCR amplification with primers based on conserved
CC motifs (see also W03426-27) of RTKs. An amplified fragment was used
CC to screen an 11.5-day-old mouse embryo and an adult mouse brain cDNA
CC library to obtain the MDK1 clone. 4 Sequence variants (see also
CC T32961-62 and W03422-25) of MDK1 were also identified. MDK1 nucleic
CC acids can be used for the recombinant prodn. of MDK1, as probes to
CC detect MDK1, and for the gene therapy of diseases involving
CC abnormalities in signal transduction, such as neurodegenerative and
CC neuroproliferative disorders and cancer.
SQ Sequence 4304 BP; 1297 A; 875 C; 1043 G; 1089 T;

```
Qy    8 tactggactcgaaacacacaaaattcaagggagctgggctggcgctggatctctattccatggg  66
          |||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db  336 TACTGGACTCGAAAGCACACAAACAGAATTGGAATGGAATTTCCTCTCCACCAGTGGG 394
```

Query Match 40.9%; Score 27; DB 1; Length 4304;
Best Local Similarity 66.1%;
Matches 39; Conservative 0; Pred. No. 0.27; Mismatches 20; Indels 0; Gaps 0;

RESULT 11
ID V58192 standard; cDNA; 4322 BP.
AC V58192;
DT 25-NOV-1998 (first entry)
DE Mouse Bsk receptor-like tyrosine kinase cDNA clone.
KW Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW neurodegenerative disease; limbic system neuron regeneration;
KW chromosomal abnormality; degenerative growth; development disorder;
KW viral infection; bacterial infection; Alzheimer's disease; epilepsy;
KW schizophrenia; stroke; cerebral ischaemia; ds
OS Mus sp.

FH	Key	Location/Qualifiers
FT	CDS	418..3051
FT		/*tag= a
FT		/product= "Bsk"
FT		/note= "receptor-like tyrosine kinase"

PN US581449-A.
PD 29-SEP-1998;
PF 11-JUN-1996; 673789.
PR 04-JAN-1994; US-17812.

PR 11-JUN-1996; US-673789.
PA (KROM/) KROMER L F.
PA (SCHU/) SCHULZ N T.
PA (WOUD/) WOUDE G F V.
PA (ZHOU/) ZHOU R.
PI Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR WPI; 98-541751/46.
DR P-PSDB: W71628.
PT Isolated nucleic acid sequence encoding protein - used in Bsk
PT nucleic acid probes, used in detecting alterations in level of Bsk
PT messenger-RNA in biological samples isolated from mammal afflicted
PT with disease
CC Claim 2; Fig 2; 72pp; English.
CC The present sequence encodes mouse Bsk, which is a receptor-like
CC tyrosine kinase. The nucleic acid sequence encoding Bsk can be used in
CC Bsk nucleic acid probes, which can be used in detecting alterations in
CC the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC from a mammal afflicted with a disease, such as neurodegenerative
CC diseases or disorders and neoplasms. The nucleic acid sequence can also
CC be delivered into the limbic system of patients with limbic system
CC neurodegenerative disease, disorder or injury, to promote or enhance
CC limbic system neuron regeneration or growth. Such neurodegenerative
CC diseases include, chromosomal abnormalities, degenerative growth and
CC development disorders, viral infections, Alzheimer's disease, brain
CC injuries, neoplastic conditions, bacterial infections, brain
CC schizophrenia, or stroke and cerebral ischaemia.
SQ Sequence 4322 BP; 1260 A; 887 C; 1055 G; 1120 T;

```
Qy    1 gtcaatctactggattcaaaacacaaaattcaagggagctgggctggcgctggatctctcatca  60
Db  604 GTGAATTGTTGGATTCGCGCACTGTCATGGGGACCTTGGATGGAATGCTTTTTCCAAAG 663

Qy   61 catggg  66
          |||
Db  664 AACGGG 669
```

Query Match 39.4%; Score 26; DB 1; Length 4322;
Best Local Similarity 62.1%;
Matches 41; Conservative 0; Pred. No. 0.67; Mismatches 25; Indels 0; Gaps 0;

RESULT 12
ID Q71303/c standard; cDNA to mRNA; 2005 BP.
AC Q71303;
DT 11-APR-1995 (first entry)
DE Mouse osteoblast-specific factor-6 cDNA.
KW osteoblast-specific factor; OSF-6; mouse; murine;
KW bone metabolic disease; diagnosis; therapy; ds
OS Mus musculus.

FH	Key	Location/Qualifiers
FT	cds	22..1350
FT		/*tag= a
FT		/product= OSF-6

PN EP-614974-A.
PD 14-SEP-1994.
PF 28-FEB-1994; 102992.
PR 10-MAR-1993; JP-048875.
PA (FARH) HOECHST JAPAN LTD.
PI Amann E, Kawai S, Kikuno R, Okazaki M, Takeshita S;
DR WPI; 94-281204/35.
DR P-PSDB: R60536.
PT New osteoblast specific factor -6 - active as a transcription
PT control factor in bone formation, used for treating bone
PT metabolic diseases
CC Claim 2; Page 8-11; 18pp; English.
CC Two cDNA libraries were prepared, one from murine osteoblast-like
CC cell line MC3T3-E1 and the other from mouse hepatic tissues. Double-
CC stranded fragments of 300 bp were obtained using a random extension
CC primer and were blunt-ended. The ds linker ATOS-1/2 was ligated to
CC the MC3T3-E1 cDNA and the ds linker ATOS-4/5 was ligated to liver
CC cDNA. PCR amplification of each product was performed using ATOS-1

CC   (Q71304) or ATOS-4 (Q71305), respectively, DNA specific for
CC   MC3T3-E1 was obtained by subtraction of sequences also found in the
CC   liver library, then amplification with ATOS-1. This was a first
CC   step in the cloning procedure for isolating a MC3T3-E1 specific
CC   clone, coding for OSF-6 (see Q71303).
SQ   Sequence 2005 BP; 485 A; 534 C; 552 G; 434 T;

Query Match   39.1%; Score 25.8; DB 1; Length 2005;
Best Local Similarity   63.9%; Pred. No. 0.65;
Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   6 tctactggattcaaaaacaattcaagggagctggggctggatctcttatccatcacatgg 65
          ||||| |||||||||||||||| |||||||||||| |||||| ||||||||||||||||
Db 1402 TCTACTGGTGATCCAACTACACAAGCACAAAAGGGGGTGGGGCTTGCCACAGACACTATCACAATG 1343

Qy   66 g 66
          |
Db 1342 G 1342

RESULT 13
T02946

ID   T02946 standard; cDNA; 2962 BP.
AC   T02946;
DT   16-APR-1996 (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK5 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK5;
KW   human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds
FT   1..2913
FT   /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
WPI; 95-373799/48.
P-PSDB; R85089.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PS   Claim 1; Page 41-45; 133pp; English.
CC   cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC   tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC   were isolated from a human foetal brain cDNA library using a directed
CC   PCR approach with primers (see T02960-61) based on conserved regions of
CC   receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC   extensive homology to the catalytic domain of chicken EPH-like
CC   receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC   EPH-like receptor. The isolated cDNAs are used for prodn. of
CC   recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC   to detect abnormalities in HEK receptor genes.
SQ   Sequence 2962 BP; 654 A; 914 C; 827 G; 567 T;

Query Match   35.2%; Score 23.2; DB 1; Length 2962;
Best Local Similarity   61.8%; Pred. No. 7.3;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy   7 ctactggattcaaaaacaattcaagggagctggggctggatctcttatccatcacatgg 66
          ||| |||||||||||||||||||||||||||||||||||||| ||||||||||| |||
Db   25 CTAATGGACTCACTACAGCCACTGCTGTGAGCTGGGTGTGCAATCCTCCATCAGGG 84

RESULT 14
T02948

ID   T02948 standard; cDNA; 3116 BP.
AC   T02948;
DT   16-APR-1996 (first entry)

DE   EPH-like receptor protein tyrosine kinase HEK8 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK8;
KW   human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds
FT   34..2994
FT   /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
WPI; 95-373799/48.
P-PSDB; R85091.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 1; Page 57-62; 133pp; English.
CC   cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC   tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC   were isolated from a human foetal brain cDNA library using a directed
CC   PCR approach with primers (see T02960-61) based on conserved regions of
CC   receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC   extensive homology to the catalytic domain of chicken EPH-like
CC   receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC   EPH-like receptor. The isolated cDNAs are used for prodn. of
CC   recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC   to detect abnormalities in HEK receptor genes.
SQ   Sequence 3116 BP; 720 A; 859 C; 812 G; 725 T;

Query Match   34.8%; Score 23; DB 1; Length 3116;
Best Local Similarity   68.1%; Pred. No. 8.8;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy   1 gtcaattcactggattcaaaaacaattcaagggagctggggctggat 47
          || ||| |||||||||||||| || |||||||||||||||| ||||||
Db   124 GTTACCTTATTGGATTCCAGATCTGTTCAGGGAGGAGAACTTGGGTGGGTGGAT 170

RESULT 15
T03100

ID   T03100 standard; DNA; 3348 BP.
AC   T03100;
DT   14-FEB-1996 (first entry)
DE   Protein tyrosine-kinase bpTK7 gene.
KW   Protein tyrosine-kinase; PTK; bpTK7; agonist; cell growth;
KW   differentiation; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds
FT   1..2961
FT   /*tag= a
FT   signal_peptide
FT   1..57
FT   /*tag= b
FT   mat_peptide
FT   58..2958
FT   /*tag= c
PN   WO9527061-A1.
PD   12-OCT-1995.
PF   04-APR-1995; U04228.
PR   04-APR-1994; US-222616.
PA   (GETH-) GENENTECH INC.
PI   Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI   Wood WI;
WPI; 95-366160/47.
P-PSDB; R85936.
PT   Agonist antibodies which activate specific protein tyrosine
PT   kinase(s) - also activate chimeric proteins of kinase extracellular
PT   domain and Ig constant domain, useful for studying, and therapeutic
PT   modulation of, cell growth and differentiation
PS   Disclosure; Page 88-92; 125pp; English.
CC   DNA probes based on protein tyrosine-kinase (pTK) sequences were used

us-09-104-340-7.rng

```
CC      to screen cDNA libraries to identify novel pTK genes.  The bpTKs,
CC      bpTK1, bpTK2, bpTK3, bpTK4, bpTK5 and bpTK7 (R85924-28 and R85935,
CC      respectively) are expressed in human brain tissue and show homology
CC      to known pTKs.  A full-length sequence for the bpTK7 gene (T03100)
CC      was obtd.  This gene may be used to design new drugs, peptides and
CC      antisense constructs that modulate pTK activity.
SQ      Sequence   3348 BP;   924 A;   769 C;   855 G;   800 T;


Query Match             34.8%;  Score 23;  DB 1;  Length 3348;
Best Local Similarity   68.1%;
Matches    32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         1 gtcaatctactggattcaaacaattcaagggggagctgggctggat 47
             || |  || ||||||||| | |  |||| || || ||| || |||||
Db        91 GTTACCTTATTGGATTCCAGATCTGTTCAGGGAGGAACTTGGGTGGAT 137


Search completed: May 15, 2000, 12:08:32
Job time: 4591 sec
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

```
                                    GenCore version 4.5
                        Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:           May 15, 2000, 11:43:02 ; Search time 1225.64 Seconds
                                                     (without alignments)
                                            -52.384 Million cell updates/sec

Title:            US-09-104-340-7
Perfect score:    66
Sequence:         1 gtcaactactggattcaaa...........tctcttatccatcacatggg 66

Scoring table:    IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters:       1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :  GenEmbl:*
            1: gb_ba1:*            45: gb_htg7:*
            2: gb_ba2:*            46: em_htg1:*
            3: gb_om:*             47: em_htg2:*
            4: gb_ov:*             48: em_htg3:*
            5: gb_pat:*            49: em_hum5:*
            6: gb_ph:*             50: gb_pl3:*
            7: gb_pl1:*            51: gb_pr5:*
            8: gb_pl2:*            52: gb_htg8:*
            9: gb_pr1:*            53: gb_htg9:*
           10: gb_pr2:*            54: gb_htg10:*
           11: gb_pr3:*            55: gb_htg11:*
           12: gb_ro:*             56: gb_htg12:*
           13: gb_sts:*            57: gb_htg13:*
           14: gb_sy:*             58: gb_htg14:*
           15: gb_un:*
           16: gb_vi:*
           17: em_fun:*
           18: em_hum1:*
           19: em_hum2:*
           20: em_in:*
           21: em_om:*
           22: em_or:*
           23: em_ov:*
           24: em_pat:*
           25: em_ph:*
           26: em_pl:*
           27: em_pr:*
           28: em_ro:*
           29: em_sts:*
           30: em_sy:*
           31: em_un:*
           32: em_vi:*
           33: gb_htg1:*
           34: gb_htg2:*
           35: gb_in1:*
           36: gb_in2:*
           37: em_ba1:*
           38: em_ba2:*
           39: em_hum3:*
           40: em_hum4:*
           41: gb_pr4:*
           42: gb_htg3:*
           43: gb_htg4:*
           44: gb_htg5:*
           45: gb_htg6:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Pred	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	66	100.0	3132	5	A28003	A28003 H.sapiens H
	2	66	100.0	3132	5	I68018	I68018 Sequence 9
	3	66	100.0	3149	9	HUMHEK	M83941 Human recep
	4	59.6	90.3	3441	5	CHCCEK4	M68514 Chicken eph
	5	59.6	90.3	3254	5	I15007	I15007 Sequence 15
	6	54.8	83.0	2032	12	MUSMEK4SE	M68515 Mouse eph-r
	7	54.8	83.0	3197	12	MUSMEK4	M68513 Mouse eph-r
	8	53.2	80.6	3077	12	RND69278	U69278 Rattus norv
c	9	34	51.5	3162	10	HUMRPTKB	L36644 Homo sapien
	10	34	51.5	3903	10	HSEHK1	X95425 H.sapiens m
	11	32.4	49.1	4124	5	GGCEK7B	U03910 Gallus gall
	12	32.4	49.1	4737	12	MMU72207	U72207 Mus musculu
c	13	32.4	49.1	160703	10	HS61A9	AL035703 Human DNA
	14	31.2	47.3	4523	10	HUMRPTK	L36642 Homo sapien
	15	31.2	47.3	116490	32	HSJ1889K14	AL121966 Homo sapi
	16	31.2	47.3	143140	43	AC015609	AC015609 Homo sapi
c	17	30	45.5	184277	55	AC021955	AC021955 Homo sapi
	18	29.8	45.2	3042	4	XLSEK1	X91191 X.laevis mR
	19	29.8	45.2	3193	12	XELPAGAAA	L26099 Xenopus lae
	20	29.2	44.2	3531	12	RNEHK1	X78689 R.norvegicu
	21	29.2	44.2	3906	5	AR062744	AR062744 Sequence
	22	29.2	44.2	4165	41	AR025488	AR025488 Sequence
c	23	28.4	43.0	132805	41	AC009425	AC009425 Homo sapi
	24	27.6	41.8	139334	33	GGY14271	Y14271 Gallus gall
	25	27.6	41.8	3943	4	ALI38846	AL138846 Homo sapi
	26	27	40.9	2006	12	RNU21955	U21955 Rattus norv
	27	27	40.9	2323	12	MMKIN1T2	X79084 M.musculus
	28	27	40.9	2901	12	MMKIN1T1	X79083 M.musculus
	29	27	40.9	3208	12	RNU21954	U21954 Rattus norv
	30	27	40.9	3943	12	MMU58332	U58332 Mus musculu
	31	27	40.9	4304	12	MMKIN1	X79082 M.musculus
c	32	26.8	40.6	105210	40	AC006383	AC006383 Homo sapi
c	33	26.6	40.3	170804	40	AC009411	AC009411 Homo sapi
	34	26.6	40.3	185587	43	AC009496	AC009496 Homo sapi
c	35	26.4	40.0	4664	4	XLA2493	XLA2493 Xenopus l
c	36	26.2	39.7	4242	12	MMSEK	X65138 M.musculus
	37	26.2	39.7	4242	42	S57168	S57168 Sek-Eph-rel
	38	26.2	39.7	77484	42	AC015510	AC015510 Homo sapi
c	39	26.2	39.7	170504	33	HSD255D2	AL109928 Homo sapi
c	40	26	39.4	3600	4	CHKCEK8	D38174 Gallus gall
	41	26	39.4	4322	5	AR043381	AR043381 Sequence
	42	26	39.4	4322	12	MMU07357	U07357 Mus musculu
	43	26	39.4	4577	4	DRU89295	U89295 Danio rerio
	44	25.8	39.1	1988	12	MMU57413	U57413 Mus musculu
	45	25.8	39.1	2000	12	MMU40930	U40930 Mus musculu

ALIGNMENTS

RESULT 1

LOCUS　　　A28003　　　　3132 bp　　DNA　　　　　　PAT　　25-SEP-1995
DEFINITION　H.sapiens HEK gene.
ACCESSION　A28003
VERSION　　A28003.1　GI:1247486
KEYWORDS　　.
SOURCE　　　human.
ORGANISM　　Homo sapiens
　　　　　　Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
　　　　　　Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1 (bases 1 to 3132)
AUTHORS
TITLE　　　A NOVEL RECEPTOR-TYPE TYROSINE KINASE AND USE THEREOF
JOURNAL　　Patent: WO 9300425-A 7 07-JAN-1993;
FEATURES　　　　　　　　　Location/Qualifiers
source　　　1..3132
　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　/db_xref="taxon:9606"
gene　　　　100..3051
　　　　　　　　　/gene="HEK"
CDS　　　　　100..3051
　　　　　　　　　/gene="HEK"
　　　　　　　　　/codon_start=1
　　　　　　　　　/protein_id="CAA01906.1"
　　　　　　　　　/db_xref="GI:1247487"
　　　　　　　　　/db_xref="SWISS-PROT:P29320"
　　　　　　　　　/translation="MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGW
　　　　　　　　　ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
　　　　　　　　　LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDR
　　　　　　　　　ILKLNTEIREYGPVNKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
　　　　　　　　　DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
　　　　　　　　　FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKDPPSMACTRPPSSPRNVIS
　　　　　　　　　NINETSVILDWSWPLDTGGRKDVTFNICKKCGWNIKQCEPCSPNVRFLPRQFGLTNT
　　　　　　　　　TVTVDLLAHYNYTFEIDAVNGVSELSSPPRQFAAVSITNQAAPSPVLTIKKDRTSR
　　　　　　　　　NSISLSWQEPEHPNGIILDYEVKYKEQOETSYTILRARGTNVTISSLKPDTIYVLQ
　　　　　　　　　IRARTAAGYGTNSRKPEFETSPDSFSISGESSQVVMIAISAAVAILLTVVIYVLIGR
　　　　　　　　　FCGYKSKHGADEKRLHFGNGHLKLPGLRTYVDPHTYEDPTQAVHEFAKELDATNISID
　　　　　　　　　KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
　　　　　　　　　IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLLGMLRGIASGMKYLSDM
　　　　　　　　　GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
　　　　　　　　　RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
　　　　　　　　　MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDISTFR
　　　　　　　　　TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
BASE COUNT　　888 a　709 c　761 g　774 t
ORIGIN

RESULT 2

　　　　　　Query Match　　　100.0%;　Score 66;　DB 5;　Length 3132;
　　　　　　Best Local Similarity　100.0%;　Pred. No. 2.3e-15;
　　　　　　Matches　66;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　　1 gtcaattcactggattcaaaaacaattcaaggggagctgggatctcttatccatca 60
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　187 GTCAATTCACTGGATTCAAAAACAATTCAAGGGGAGCTGGGATCTCTTATCCATCA 246

Qy　　61 catggg 66
　　　　　 ||||||
Db　247 CATGGG 252

LOCUS　　I68018　　　　3132 bp　　DNA　　　　　　PAT　　04-FEB-1998
DEFINITION　Sequence 9 from patent US 5674691.
ACCESSION　I68018
VERSION　　I68018.1　GI:2830140
KEYWORDS　　.
SOURCE　　　Unknown.
ORGANISM　　Unknown.
　　　　　　Unclassified.

REFERENCE　1 (bases 1 to 3132)
AUTHORS　　Boyd,A.W., Simpson,R.John, Wicks,I., Ward,L.David and Wilkinson,D.
TITLE　　　Method of screening for ligands to a receptor-type tyrosine kinase
JOURNAL　　Patent: US 5674691-A 9 07-OCT-1997;
FEATURES　　　　　　　　　Location/Qualifiers
source　　　1..3132
　　　　　　　　　/organism="unknown"
BASE COUNT　　888 a　709 c　761 g　774 t
ORIGIN

RESULT 3

　　　　　　Query Match　　　100.0%;　Score 66;　DB 5;　Length 3132;
　　　　　　Best Local Similarity　100.0%;　Pred. No. 2.3e-15;
　　　　　　Matches　66;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　　1 gtcaattcactggattcaaaaacaattcaaggggagctgggatctcttatccatca 60
　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　187 GTCAATTCACTGGATTCAAAAACAATTCAAGGGGAGCTGGGATCTCTTATCCATCA 246

Qy　　61 catggg 66
　　　　　 ||||||
Db　247 CATGGG 252

LOCUS　　HUMHEK　　　　3149 bp　　mRNA　　　　　PRI　　31-DEC-1994
DEFINITION　Human receptor tyrosine kinase (HEK) mRNA, complete cds.
ACCESSION　M83941
VERSION　　M83941.1　GI:183931
KEYWORDS　　receptor protein-tyrosine kinase.
SOURCE　　　Homo sapiens lymphoid tumor cDNA to mRNA.
ORGANISM　　Homo sapiens
　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
　　　　　　Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1 (bases 1 to 3149)
AUTHORS　　Wicks,I.P., Wilkinson,D., Salvaris,E. and Boyd,A.W.
TITLE　　　Molecular cloning of HEK, the gene encoding a receptor tyrosine
　　　　　　kinase expressed by human lymphoid tumor cell lines
JOURNAL　　Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1611-1615 (1992)
MEDLINE　　92179233
FEATURES　　　　　　　　　Location/Qualifiers
source　　　1..3149
　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　/cell_line="LK63"
　　　　　　　　　/tissue_type="lymphoid tumor"
gene　　　　101..3052
　　　　　　　　　/gene="HEK"
CDS　　　　　101..3052
　　　　　　　　　/gene="HEK"
　　　　　　　　　/codon_start=1
　　　　　　　　　/product="receptor protein kinase"
　　　　　　　　　/db_xref="GI:183932"
　　　　　　　　　/translation="MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKIIQGELGW
　　　　　　　　　ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFT
　　　　　　　　　LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDR
　　　　　　　　　ILKLNTEIREYGPVNKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
　　　　　　　　　DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
　　　　　　　　　FYKALDGNMKCAKCPPHSSTQEDGSMNCRCENNYFRADKDPPSMACTRPPSSPRNVIS
　　　　　　　　　NINETSVILDWSWPLDTGGRKDVTFNICKKCGWNIKQCEPCSPNVRFLPRQFGLTNT
　　　　　　　　　TVTVDLLAHYNYTFEIDAVNGVSELSSPPRQFAAVSITNQAAPSPVLTIKKDRTSR
　　　　　　　　　NSISLSWQEPEHPNGIILDYEVKYKEQOETSYTILRARGTNVTISSLKPDTIYVFQ
　　　　　　　　　IRARTAAGYGTNSRKPEFETSPDSFSISGESSQVVMIAISAAVAILLTVVIYVLIGR
　　　　　　　　　FCGYKSKHGADEKRLHFGNGHLKLPGLRTYVDPHTYEDPTQAVHEFAKELDATNISID
　　　　　　　　　KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
　　　　　　　　　IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLLGMLRGIASGMKYLSDM
　　　　　　　　　GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
　　　　　　　　　RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
　　　　　　　　　MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDISTFR
　　　　　　　　　TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
　　　　　　　　　ETQSKNGPVPV"

```
BASE COUNT
ORIGIN
          891 a    711 c    768 g    779 t

Query Match    100.0%;    Score 66;    DB 9;    Length 3149;
Best Local Similarity 100.0%;    Pred. No. 2.3e-15;
Matches 66;    Conservative 0;    Mismatches 0;    Indels 0;    Gaps 0;

Qy    1 gtcaatctactggattcaaaaacaaattcaaggggagactgggcctggggatctcttatccatca 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 GTCAATCTACTGGATTCAAAAACAAATTCAAGGGGAGAGCTGGGCCTGGGGCTGGGGATCTCTTATCCATCA 247

Qy   61 catggg 66
        ||||||
Db  248 CATGGG 253

RESULT 4
CHKCEK4
LOCUS       CHKCEK4    3241 bp    mRNA    VRT    15-FEB-1994
DEFINITION  Chicken eph-related receptor tyrosine kinase (Cek4) mRNA, complete
            cds.
ACCESSION   M68514
VERSION     M68514.1 GI:454809
KEYWORDS    receptor tyrosine kinase.
SOURCE      Gallus gallus
ORGANISM    Gallus gallus 10 day embryo cDNA to mRNA.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 3241)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
COMMENT     On Feb 16, 1994 this sequence version replaced gi:211446.
FEATURES             Location/Qualifiers
     source          1..3241
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /dev_stage="10 day embryo"
                     /tissue_type="embryo"
     gene            32..2983
                     /gene="Cek4"
     CDS             32..2983
                     /gene="Cek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA48666.1"
                     /db_xref="GI:211447"
                     /translation="MDRRLPLLLLCAALGSAGRLSARPGNEVNLLDSKTIQGELGWI
                     SYPSHGWEEISGVDEHYTPIRYQESNVMDHSQNNWLRTNWIPRNSAQKIVELKFTL
                     RDCNSIPLVLGTCKETFNLYYMESDDHLAKFREHQFTKIDTIAADESFTQMDLGDRI
                     LKINTEVREVGPVSKGFYILAFQDVGACVALVSVRVFYKKCPSVVKNLAMFPDTVPMD
                     SQSLVEVRGSCVNHSKEEPPKMYCSTEGEWLVPIGKCLCNAGYEERGFACQACRPGF
                     YKASAGNVKCAKCPPHSSTYEDASLNCRCEKNYFRSEKDPPSMACTRPPSAPRNVISN
                     INETSVILLDWSWPLDTGGRKDVTFNICKKCGGSSKICEPCSDNVRFLPRQTGLNWTT
                     VTVVDLLAHHNYTFEIDAVNGVSDLSTLSRQFAAVSITTNQAAPSPITVIRKDRISRN
                     SVSLSWQEPEHPNGIILDYEVKYYEKQEETSYTILRAKSTNVTISGLKPDTTYVFQI
                     RARTAARYGTSSSRKFEFETSPDSFSISSENSQVVMIAISAAVAILLLTVVVVLGRF
                     CGYKSKHGTDEKRLHFGNGHLKLPGLRTIYVDPHTYEDPNQAVHEFAKELDASNISID
                     KVGAGEFGEVCSGRLKLPGKREISVAIKTLKAGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVVTIKSRPVMIYTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILVNSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDAWSYGIVLWEVMSGERPWEMSPQDVIKAVDEGYRLPPPMDCPAALTYQL
                     MLDCWQKDRNRNRPKFEQIVSILDKLIRNPSSLKIITNAAARPSNLLLDQSNIDIAFR
                     TAGDWLNGFRGQCKGIFTGVEYSSCDTIAKISTDDMKRGVTVVGPQKKIVSSIKTL
                     ETHTKNSPVPV"
BASE COUNT     916 a    735 c    794 g    796 t
ORIGIN
     polyA_site      3241


Query Match    90.3%;    Score 59.6;    DB 4;    Length 3241;
Best Local Similarity 93.9%;    Pred. No. 7.3e-13;
Matches 62;    Conservative 0;    Mismatches 4;    Indels 0;    Gaps 0;

Qy    1 gtcaatctactggattcaaaaacaaattcaaggggagactgggcctggggatctcttatccatca 60
        || ||||| ||||||||||||||||||||||||||| ||||||||||||||||| || ||||| ||
Db  116 GTTAATCTCGCTGGATTCAAAAACAAATTCAAAGGGGAGAGCTGGGCCTGGATCTCCTACCCATCA 175

Qy   61 catggg 66
        ||||||
Db  176 CATGGG 181

RESULT 5
I15007
LOCUS       I15007    3254 bp    DNA    PAT    02-APR-1996
DEFINITION  Sequence 15 from patent US 5457048.
ACCESSION   I15007
VERSION     I15007.1 GI:1249915
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 3254)
  AUTHORS   Pasquale,E.B. and Sajjadi,F.G.
  TITLE     Eph-related tyrosine kinases, nucleotide sequences and methods of
            use
  JOURNAL   Patent: US 5457048-A 15 10-OCT-1995;
FEATURES             Location/Qualifiers
     source          1..3254
                     /organism="unknown"
BASE COUNT     926 a    737 c    796 g    795 t
ORIGIN


Query Match    90.3%;    Score 59.6;    DB 5;    Length 3254;
Best Local Similarity 93.9%;    Pred. No. 7.3e-13;
Matches 62;    Conservative 0;    Mismatches 4;    Indels 0;    Gaps 0;

Qy    1 gtcaatctactggattcaaaaacaaattcaaggggagactgggcctggggatctcttatccatca 60
        || ||||| ||||||||||||||||||||||||||| ||||||||||||||||| || ||||| ||
Db  116 GTTAATCTCGCTGGATTCAAAAACAAATTCAAAGGGGAGAGCTGGGCCTGGATCTCCTACCCATCA 175

Qy   61 catggg 66
        ||||||
Db  176 CATGGG 181

RESULT 6
MUSMEK4SE
LOCUS       MUSMEK4SE    2032 bp    mRNA    ROD    15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (Mek4) secreted mRNA,
            complete cds.
ACCESSION   M68515
VERSION     M68515.1 GI:454828
KEYWORDS    receptor tyrosine kinase.
SOURCE      Mus musculus
ORGANISM    Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
            cDNA to mRNA.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2032)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
COMMENT     On Feb 16, 1994 this sequence version replaced gi:199121.
FEATURES             Location/Qualifiers
     source          1..2032
                     /organism="Mus musculus"
```

RESULT 7 (left column)

```
CDS             /strain="IRC x Swiss Webster"
                /db_xref="taxon:10090"
                /dev_stage="11.5 day embryo"
                /tissue_type="embryo"
                89. .1702
                /codon_start=1
                /product="Mek4 secreted"
                /protein_id="AAA39522.1"
                /db_xref="GI:454829"
                /translation="MDCHLSILVLLGCCVLSSGGELSPQPSNEVNLLDSKTIQGELGW
                ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                LRDCNSIPLVLGCKETFNLYLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRI
                LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMD
                SQSLVEVRGSCVNNSKEEDPPRMCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                INETSVILDWSWPLDTGGRKDIIFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                VTVTDLLAHTNYTFEIDAVNGVSELSSPPRQYAAVSITTNQAAPSPVMTIKKDRTSRN
                SISLSWQEPEHPNGIILDYEVKYYQKQEQETSYTILRARGTNVTISSLKPDTIYVFQI
                RARTAAGYGTNSRKFEFETSPDCMYFSF"
        polyA_signal    2010. .2015
        polyA_site      2032
BASE COUNT        578 a    473 c    460 g    521 t
ORIGIN
Qy    1  gtcaatctactggattcaaaaaacaaattcaaggggagctgggctggatctccatca  60
         ||||  ||  |||| |||||  ||| |||||||| ||||||||||||||||| ||||||
Db  176  GTTAATCTACTACTAGATTCGAAAACAATTCAAGGAGAGCTGGGCTGGATCTCCTACCATCC  235

Qy   61  catggg  66
         ||||||
Db  236  CATGGG  241
```

Query Match   83.0%;   Score 54.8;   DB 12;   Length 2032;
Best Local Similarity   89.4%;
Matches 59;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

```
LOCUS       MUSMEK4      3197 bp    mRNA            ROD       15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (MeK4) mRNA, complete
            cds.
ACCESSION   M68513
VERSION     M68513.1  GI:199119
KEYWORDS    .
SOURCE      Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
            cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3197)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92031278
FEATURES             Location/Qualifiers
     source          1. .3197
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     gene            89. .3040
                     /gene="Mek4"
     CDS             89. .3040
                     /gene="Mek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA39521.1"
                     /db_xref="GI:199120"
```

RESULT 8 (right column)

```
BASE COUNT        907 a    767 c    775 g    748 t
ORIGIN
                /translation="MDCHLSILVLLGCCVLSCSGGELSPQPSNEVNLLDSKTIQGELGW
                ISYPSHGWEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                LRDCNSIPLVLGCKETFNLYLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRI
                LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMD
                SQSLVEVRGSCVNNSKEEDPPRMCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                INETSVILDWSWPLDTGGRKDIIFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNTT
                VTVTDLLAHTNYTFEIDAVNGVSELSSPPRQYAAVSITTNQAAPSPVMTIKKDRTSRN
                SISLSWQEPEHPNGIILDYEVKYYQKQEQETSYTILRARGTNVTISSLKPDTIYVFQI
                RARTAAGYGTNSRKFEFETSPDCFSFSEGSENSHVVMIAISAAVAIIVLTVVLVGRF
                CGYHKSKHSAEKRLHFGNGHLKLPGLRTIVDPHTYEDPTQAVHEFAKELDANVSID
                KVGGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPNI
                IRLEGVVTRKSPVMIITEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                GYVHRDLAARNILVNSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAMSY
                RKFTSASDVWSYGIVLWEVMSYGERPYSQMSNQDVIKAVDERYRLPPPMDCPAALYQL
                MLDCWQKDRNRRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDIATFH
                TTGDWLNGMRTAHCKEIFTGVEYSSCDIIAKISTDDMRKVGVTVVGPQKKIISTKAL
                ETQSKNGPVPV"
BASE COUNT
Qy    1  gtcaatctactggattcaaaaaacaaattcaaggggagctgggctggatctccatca  60
         ||||  ||  |||| |||||  ||| |||||||| ||||||||||||||||| ||||||
Db  176  GTTAATCTACTACTAGATTCGAAAACAATTCAAGGAGAGCTGGGCTGGATCTCCTACCATCC  235

Qy   61  catggg  66
         ||||||
Db  236  CATGGG  241
```

Query Match   83.0%;   Score 54.8;   DB 12;   Length 3197;
Best Local Similarity   89.4%;
Matches 59;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

```
LOCUS       RNU69278     3077 bp    mRNA            ROD       23-MAR-1998
DEFINITION  Rattus norvegicus eph-related receptor tyrosine kinase homolog
            (ReK4) mRNA, complete cds.
ACCESSION   U69278
VERSION     U69278.1  GI:1943913
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McIlernan,C.F. and Feldman,A.M.
  TITLE     IL-1 beta alters the expression of the receptor tyrosine kinase
            gene r-EphA3 in neonatal rat cardiomyocytes
  JOURNAL   Am. J. Physiol. 274 (1), H331-H341 (1998)
  MEDLINE   98120505
REFERENCE   2  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McIlernan,C.F. and Feldman,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1996) Cardiology, University of Pittsburgh,
            Lothrop Street, Pittsburgh, PA 15213, USA
COMMENT     On Apr 18, 1997 this sequence version replaced gi:1698721.
FEATURES             Location/Qualifiers
     source          1. .3077
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /dev_stage="neonatal"
                     /cell_type="cardiomyocytes"
     gene            1. .3077
                     /gene="Rek4"
     CDS             35. .2989
                     /gene="Rek4"
                     /codon_start=1
                     /product="eph-related receptor tyrosine kinase homolog"
                     /protein_id="AAC06273.1"
```

```
                    /db_xref="GI:1698722"
                    /translation="MDCHLSILILFGCCVLSCSRELSPQPSNEVNLLDSKTIQGELGW
                    ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                    LRDCNSLPLVLGTCKETFNLYYMESDDDGHGVKFREQFKTKIDIIAADESFTQMDLGDR
                    ILKLNTEIREVGPVNKGFYLAFQDVGACVALVSVRVFKKCPTVKNLAMFPDTVPM
                    DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPG
                    FYKALDGVAKCTCHPHSSTQEDGSMNCRCENNYFRAEKDPPSMACTRPPSAPRNVIS
                    NINETSVILDWSWPLDTGGRKDITFNICKKCGWNVRQCEPCSPNVRFLPRQLGLTNT
                    TVTVDLLAHTNYTFEIDAINGVSELSSPPRQFAAVSITTNQAAPSPVMTIKKDRTSR
                    NSISLSWQEPEHPNGIILDYEVKYYEKQEETSYTILRARGINVTISSLKPDTTYVFQ
                    IRARTAACYGTNSRKFEFETNNPSDSFSIGGENSHVVMIAISAAVIIVLTVTIVLVGR
                    FCGYHKSHSSDEKRLHFGNGHLRLPGLRTIVDPHTYEDPTQAVHEFAKELDATNIAI
                    DKVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPN
                    IIRLEGVVTKSRPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVLGMLRGIASGMKYLSD
                    MGYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPVRWTSPEATA
                    YRKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPLPMDCPAALYQ
                    LMLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDIATF
                    HTTGDWLNGMRTAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKA
                    LETQSKNGPVPV"
```

```
BASE COUNT     877 a     714 c     758 g     728 t
ORIGIN
```

```
Qy          1 gtcaatctactggattcaaaaacaattcaaggggagctgggctggatctcttatccatca 60
              | ||||||||||||||||||||||||||||||||| ||||||||||| |||||||||||
Db        122 GTTAATCTACTAGATTCAAAAACGAATTCAAGGAGAGAGCTGGGCTGGATCTCCTACCCATCC 181
```

```
Qy         61 catggg 66
              || |||
Db        182 CACGGG 187
```

```
RESULT  9
LOCUS       HUMRPTKB     3162 bp     mRNA                    PRI      10-AUG-1995
DEFINITION  Homo sapiens receptor protein-tyrosine kinase (HEK7) mRNA, 3' end.
ACCESSION   L36644
VERSION     L36644.1     GI:551611
KEYWORDS    EPH-like receptor PTK; receptor protein-tyrosine kinase.
SOURCE      Homo sapiens (clone brain: Stratagene premade library, cat
            #936206) female fetus, 17-18 weeks gestation brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3162)
  AUTHORS   Fox,G.M., Holst,P.L., Chute,H.T., Lindberg,R.A., Janssen,A.M.,
            Basu,R. and Welcher,A.A.
  TITLE     cDNA cloning and tissue distribution of five human EPH-like
            receptor protein-tyrosine kinases
  JOURNAL   Oncogene 10 (5), 897-905 (1995)
  MEDLINE   95206782
FEATURES             Location/Qualifiers
     source          1..3162
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /sex="female"
                     /dev_stage="fetus, 17-18 weeks gestation"
                     /clone_lib="Stratagene premade library, cat #936206"
                     /tissue_type="brain".
     mRNA            <1..3162
                     /gene="HEK7"
                     /product="receptor protein-tyrosine kinase"
     gene            1..3162
                     /gene="HEK7"
     CDS             <1..2976
                     /gene="HEK7"
                     /codon_start=1
                     /product="receptor protein-tyrosine kinase"
                     /protein_id="AAA74245.1"
```

```
Query Match    Best Local Similarity    80.6%;    Score 53.2;  DB 12;  Length 3077;
Matches   58;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;
```

```
                    /db_xref="GI:551612"
                    /translation="PASLAGCYSAPRRAPLWTCLILCAALRTLLASPSNEVNLLDSRT
                    VMGDLGWIAPFKNGWEEIGEVDENAPAIHYTQVCKVMEQNNWLLTSWIGSNEGASRI
                    FIELKFTLRDCNSLPGLGTCKETFNLYYMESDDDGNGRNIKENQYIKIDIIAADESFT
                    ELDLGDRVMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVYKKCPSVVRHLAV
                    FPDITGADSSQLLEVSGSCVNHSVTDEPPKMHCSAEGEWLVPIGKCMCKAGYEEKNG
                    TCQVCRPGFFKASPHIQSGSGCPHSYTHEEAASTSCVCEKDYFRESDPPTMACTRPP
                    SAPRNAISNVNETSVFLEWIPPADTGGRKDVSYIIACKKCNSHAGVCEEGGHVRLP
                    RQSGLKNTSVMMVDLLAHTNYTFEIEAVNGVSDLSPGARQYVSVNVTTNQAAPSPVTN
                    VKKGRKNSISLSWQEPDRPNGIILEYEIKYFEKDQETSYTILRARGVSRYLKSETTITAEGLKP
                    ASVVFQIRARTAAGYGVFSRRFEFETTPVFAAASSDOSQIPVIAVSVTVGLLAVI
                    GVLLSGRRCGYSKAKQDPEEEKMHFHNGHIKLPGVRTIYDPHTYEDPNQAVHEFAKEI
                    EASCITIERVIGAGEFGEVCSGRLKLPGRRELPVAIKTLKVGYTEKQRRDFLGEASIM
                    GQFDHPNIIHLEGVVTKKSPVMIVTEYMENGSLDTFLKKNDGQFTVIQLVGMLRGISA
                    GMKYLSDMGYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRW
                    GVLLSGRRCGYSKAKQDPEEKMHFHNGHIKLPGVRTIYDPHTYEDPNQAVHEFAKEI
                    CPAALYQLMLDCWQKERNSRPKFDEIVNMLDKLIRNPSSLKTLVNASCRVSNLLAEHS
                    PLGSGAYRSVGEWLEAIKMGRYTEIFMENGYSSSMDAVAQVTLEDLRRLGVTLVGHQKK
                    IMNSLQEMRVQLVGNGMVPL"
```

```
BASE COUNT     922 a     667 c     774 g     799 t
ORIGIN
     3'UTR
                     /gene="HEK7"
                     2977..3162
```

```
Qy          1 gtcaatctactggattcaaaaacaattcaaggggagctgggctggatctcttatccatca 60
              | ||||||||||||||||||||||||||||||| ||||||||||||| |||||||||||
Db        109 GTGAATTAATTGGATTCACGGCACTGTCATGGGGGACCTGGGATTGCTTTTCCAAAA 168
```

```
Qy         61 catggg 66
              || |||
Db        169 AATGGG 174
```

```
RESULT  10
LOCUS       HSEHK1     3903 bp     mRNA                    PRI      11-AUG-1999
DEFINITION  H.sapiens mRNA for EHK-1 receptor tyrosine kinase.
ACCESSION   X95425
VERSION     X95425.1     GI:1177465
KEYWORDS    EHK-1; receptor tyrosine kinase.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3903)
  AUTHORS   Miescher,G.C., Taylor,V., Olivieri,G., Mindermann,T., Schrock,E.
            and Steck,A.J.
  TITLE     Extensive splice variation and localization of the EHK-1 receptor
            tyrosine kinase in adult human brain and glial tumors
  JOURNAL   Brain Res. Mol. Brain Res. 46 (1-2), 17-24 (1997)
  MEDLINE   97334377
REFERENCE   2  (bases 1 to 3903)
  AUTHORS   Miescher Constant,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1996) Guido C. Miescher, University Hospitals
            Basel, Department of Research, Departement Forschung, Kantonsspital
            Basel, 4031 Basel, Switzerland
  COMMENT   Overlaps with L36642-L36645
            Partial human EHK-1 cDNA without information on mRNA splicing
            variants has been published by Fox, G.M. et al. (1995). Oncogene
            10:897-905.
FEATURES             Location/Qualifiers
     source          1..3903
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone_lib="lambda ZAPII #936206"
                     /clone="HFB415"
```

```
Query Match    Best Local Similarity    51.5%;    Score 34;  DB 10;  Length 3162;
Matches   46;  Conservative    0;  Mismatches    20;  Indels    0;  Gaps    0;
```

RESULT 11
GGCEK7B

```
LOCUS       GGCEK7B      4124 bp    mRNA            VRT       30-NOV-1995
DEFINITION  Gallus gallus receptor-type protein-tyrosine kinase precursor
            (Cek7) mRNA, complete cds.
ACCESSION   U03910
VERSION     U03910.1  GI:555617
KEYWORDS    .
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 4124)
  AUTHORS   Siever,D.A. and Verderame,M.F.
  TITLE     Identification of a complete Cek7 receptor protein tyrosine kinase
            coding sequence and cDNAs of alternatively spliced transcripts
  JOURNAL   Gene 148 (2), 219-226 (1994)
  MEDLINE   95047429
REFERENCE   2  (bases 1 to 4124)
  AUTHORS   Siever,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1993) Doyle A. Siever, Microbiology and
            Immunology, The Pennsylvania State University College of Medicine,
            P.O. Box 850, Hershey, PA 17033, USA
FEATURES             Location/Qualifiers
     source          1..4124
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="overlapping clones: 1-13, 2-6, and 9-13"
                     /tissue_type="body wall"
                     /dev_stage="ten-day embryo"
                     /note="this sequence diverges 3' of nucleotide 3023 when
                     compared with previously reported Cek7 sequence (GenBank
                     Accession Number Z19058)"
     5'UTR           <1..84
     mRNA            join(<1..1801,1868..4124)
                     /note="alternative splicing removes nucleotides 1802-1867
                     (insertion B)"
     mRNA            <1..4124
     gene            85..3126
                     /gene="Cek7"
     CDS             join(85..921,1414..1801,1868..3126)
                     /gene="Cek7"
                     /note="product of alternatively spliced mRNA lacking
                     insertion A and B"
                     /codon_start=1
                     /product="receptor-type protein-tyrosine kinase"
                     /protein_id="AAB06614.1"
                     /db_xref="GI:555620"
                     /translation="MGLRGGGRAGGPAPGWTCLLLCAALRSLLASPGSEVNLLDSRT
                     VMGDLGWIAYPRNGWEEIGEVDENYAPIHTYQVCKVMEQNQNNWLLTSWISNEGRPAS
                     SFELKFTLRDCNSLPGGLGTCKETFNMYYFESDDEDGRNIRENQYIKIDTIAADEST
                     RDFLGEASIMGQFDHPNIIHLEGVVYKSKPVMIVTEYMENGSLDTFLKKNDGQFTVIQ
                     LVGMLRGIASGMKYLSDMGYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYT
                     TRGGKIPIRWTAPEAIAFRKFTSASDVWSYGIVMWEVMSYGERPYWEMTNQDVIKAVE
                     EGYRLPSPMDCPAALYQLMLDCWQKDRNSRPKFDEIVSMLDKIIRNPSSLKTLVNASS
                     RVSNLLVEHSPVGSGAYRSVGEWLEAIKMGRYTEIFMENGYSSMDSVAQVTLEDLRL
                     GVTLVGHQKKIMNSLQEMKVQLVNGMVPL"
     gene            85..3126
                     /gene="Cek7"
                     /codon_start=1
                     /product="receptor-type protein-tyrosine kinase"
                     /protein_id="AAB06612.1"
                     /db_xref="GI:555618"
                     /translation="MGLRGGGRAGGPAPGWTCLLLCAALRSLLASPGSEVNLLDSRT
                     VMGDLGWIAYPRNGWEEIGEVDENYAPIHTYQVCKVMEQNQNNWLLTSWISNEGRPAS
                     SFELKFTLRDCNSLPGGLGTCKETFNMYYFESDDEDGRNIRENQYIKIDTIAADEST
                     ELDLGDRVMKLNTEVRDVGPLTKKGFYLAFQDVGACIALVSRVRVYKKCPSVIRNLAR
                     FPDTITGADSSQLLEVSGVCVNNHSVDTDAPKMHCSAEGEWLVPIGKCMCKAGYEEKN
                     QTGLKNTSVMWVDLLAHTNTYFEIEAVNGVSDLQQRQFSVNVTTNQAAPSVVSS
                     VKKGKITRNSISLSWQEPDRPNGIILEYEIKYFEKDQETSYTIIKSKETAITADGLKP
                     GSAYVFQIRARTAAGYGGFSRREFETFSPVLAASSDQSQIPIIVVSVTVGVILLAVVI
                     GFLLSGSCCDHGCGWASSLRAYAPSILWRCGYSKAKQDPEEEKMHFHNGHIKLPGVR
                     TYIDPHYTEDPNQAVHEFAKEIEASCITIERVIGAGEFGEVCSGRLKLQGKREPFVAI
                     KTLKVGYTEKQRRDFLGEASIMGQFDHPNIIHLEGVVTKSKPVMIVTEYMENGSLDTF
                     LKKNDGQFTVIQLVGMLRGIASGMKYLSDMGYVHRDLAARNILINSNLVCKVSDFGLS"
```

```
sig_peptide     712..783
                /product="EHK-1 receptor tyrosine kinase"
                /map="q12"
                /lab_host="XL1-blue"
                /cell_type="CNS neurons"
                /tissue_type="brain"
                /dev_stage="embryo"
                /clone="HFB129"
                /clone="HFB115"
CDS             712..3825
                /codon_start=1
                /product="EHK-1 receptor tyrosine kinase"
                /protein_id="CAA64700.1"
                /db_xref="GI:1177466"
                /db_xref="SWISS-PROT:P54756"
                /translation="MRGSGPRGAGHRRPPGGGGDTDPITPASLAGGYSAPRRAPLWTCL
                LCAALRTLASPSNEVNLLDSRTVMGDLGWIAFPKNGWEEIGEVDENYAPIHTYQVC
                KVMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPGGLGTCKETFNMYYFESDD
                ONGRNIKENQYIKIDIIAADESFTELDLGDRVMKLNTEVRDVGPLSKGFYLAFQDVG
                ACIALVSRVRVYKKCPSVVRNLAVFFDTITGADSSQLLEVSGSCVNHSVDTDEPPKHC
                SAEGEWLVPIGKCMCKAGYEEKNGTCQVCRPGFFRASPHIQSCGKCPPHSYTHEEAST
                SCVCEKDYFRESDDPTMACTRPPSAPRNAISNVNETSVFLEWIPPADTGGRKDVSYI
                IACKCKNSHAGVCEECGGHVRYLPRQSGLKNTSVMWVDLLAHTNTYFEIEAVNGVSDL
                SPGARQYVSVNVTTNQAAPSPVTNVKKGKIAKNSISLSWQEPDRPNGIILEYEIKHFE
                KDQETSYTIIKSKETTIAAGLKPASVYVFQIRARTAAGYGGFSRREFETFSPVFAAS
                SDQSQIPVIVSVTVGVILLAVVIGVLLSGSCCECGRASSLCAVAHPIIWRCGYS
                KAKQDPEEEKMHFHNGHIKLPGVRTYIDPHYTEDPNQAVHEFAKEIEASCITIERVIG
                AGEFGEVCSGRLKLQGKREPLPVAIKTLKVGYTEKORRDFLGEASIMGQFDHPNIHLE
                GVVTKSKPVMIVTEYMENGSLDTFLKKNDGQFTVIQLVGMLRGIASGMKYLSDMGYVH
                RDLAARNILINSNLVCKVSDFGLSRVLEDDPEAATTRGGKIPIRWTAPEAIAFRKFT
                SASDVWSYGIVMWEVVSYGERPYWEMTNQDVIKAVEEGYRLPSPMDCPAALYQLMLDC
                WQKENSRPKFDEIVNMLDKIIRNPSSLKTLVNASCRVSNLLAEHSPLGSGAYRSVGE
                WLEAIKMGRYTEIFMENGYSSMDAVAQVTLEDLRLGVTLVGHQRKIMNSLQEMKVQL
                VNGMVPL"
mat_peptide     784..3822
misc_feature    1622..1777
                /note="RNA splice domain I"
misc_feature    1778..2113
                /note="RNA splice domain IIa"
misc_feature    2114..2392
                /note="RNA splice domain IIb"
misc_feature    2502..2567
                /note="RNA splice domain III"
BASE COUNT      1091 a   882 c   953 g   977 t
ORIGIN
```

```
Query Match     51.5%;   Score 34;   DB 10;   Length 3903;
Best Local Similarity   69.7%;   Pred. No. 0.0077;
Matches    46;   Conservative    0;   Mismatches   20;   Indels    0;   Gaps    0;

Qy    1 gtcaatctactggatcaaaaacaatcaaggggagctggcgtggcttttcttatccatca 60
        | |||  | |||||||||| |  | |||||||||||||||||||||||||||| |
Db  892 GTGAATTATTGGATTCACGCACTGTCATGGGGACCTGGGATGGAATTGCTTTTCCAAAA 951

Qy   61 catggg 66
        |||||
Db  952 AATGGG 957
```

RVLEDDPEAAYTTRGGKIPIRWTAPEAIAFRKFTSASDVWSYGIVMWEVMSYGERPYW
EMTNQDVIKAVEEGYRLPSPMDCPAALYQLMLDCWQKDRNSRPKFDEIYSMLDKLIRN
PSSLKTLVNASSRVSNLLVEHSPVGSGAYRSVGEWLEAIKMGRYTEIFMENGYSSMDS
VAQVTLEDRRLGVTLVGHQKKIMNSLQEMKRVQLVNGMVPL"

sig_peptide   85..177
                /gene="Cek7"
                /note="predicted signal sequence"
CDS   join(85..1801,1868..3126)
                /gene="Cek7"
                /note="product of alternatively spliced mRNA lacking
                insertion B"
                /codon_start=1
                /product="receptor-type protein-tyrosine kinase"
                /protein_id="AAB60613.1"
                /db_xref="GI:555619"
                /translation="MGLRGGGRAGGPAPGWTCLLLCAALRSLLASPGSEVNLLDSRT
                VMGDLGWIAYPKNGWEEIGEVDENYAPIHIYQVCKVMEQNONNWLLTSWISNEGRPAS
                SFELKFTLRDCNSLPGLGTGTCKETFNMYYFESDDEDGRNIRENQYIKIDTIAADESFT
                ELDGDRVMKLNTEVRDVGPLTKRGFYLAPQDVGACIALVSVRYYKRYKPSVIRNLAR
                FPDITGADSSOLLEVSGVCVNHSVTDEAPKMHCSAEGEWLVPIGKCLCKAGYEEKNN
                TCQVCRPGFFKASPHSPCSKCPPHSYTLDEASTCLCEEHYFRRESDPPTMACTRPP
                SAPRSAISNVNETSYFLEWIPPADTGGRDVSYIIACKKCNSHSGICEGACGSHYRLYP
                QQTGLKNTSVMMVDLLAHHNYTFEIEAVNGVSDOSPGARQFVSVNVTTNQAAPSPVSS
                VKRGKITKNSISLSWQEPDRPNGIILEYEIKYFEKDQETSYTIIKSKETAITADGLKP
                GSAVFOIRARTAAGYGGFSRREFEFTSPVLAASSDQSQIPIIVVSVTGVCILLAVVI
                GFLSGRRCGYSRAKQDPEEEKMHFHNGHKLPGVRTYIDPHTYEDPNQAVHERAKEI
                EASCITIERVIGAGEFGEVCSGRLKLQGKREFPVAIKTLKVGYTEAQRRDFLGEASIM
                GOFDHPNIIHLEGVVTKSKPVMIIVTEYMENGSLDTFLKNDQOFTVIQLVGMLRGIAS
                GMKYLSDMGYVHRDLAARNILNSNLVCKVSDFGLSRVLEDDPEAAYTRGGKIPIRW
                TAPEAIAFRKFTSASDVWSYGIVMWEVMSYGERPYWEMTNQDVIRAVEEGYRLPSPMD
                CPAALYQLMLDCWQRDRNSRPKFDEIVSMLDKLIRNPSSLKTLVNASSRVSNLLVEHS
                PVGSGAYRSVGEWLEAIKMGRYTEIFMENGYSSMDSVAQVTLEDLRRLGVTLVGHQKK
                IMNSLQEMKVQLVNGMVPL"
variation   642
                /gene="Cek7"
                /note="polymorphism in third position of the codon which
                does not alter the amino acid sequence"
                /replace="t"
variation   771
                /gene="Cek7"
                /note="polymorphism in third position of the codon which
                does not alter the amino acid sequence"
                /replace="a"
variation   879
                /gene="Cek7"
                /note="polymorphism in third position of the codon which
                does not alter the amino acid sequence"
                /replace="t"
variation   882
                /gene="Cek7"
                /note="polymorphism in third position of the codon which
                does not alter the amino acid sequence"
                /replace="g"
misc_feature   1726..1797
                /gene="Cek7"
                /note="hydrophobic region characteristic of a
                transmembrane domain"
3'UTR   3127..4124
polyA_site   4124
BASE COUNT   1103 a   917 c   1060 g   1044 t
ORIGIN

Qy   1 gtcaatctactggattcaaaaacaattcaagggagctgggctggatctcttatccatca 60
         |||||||||||||||||||||||||||| |||||||||||||||||| ||||| |||||
Db   193 GTGAATTACTGGACTCACCGCACACTAATGGGGACTAGGATGGATGAGCTACCACCAAAA 252

Qy   61 gtcaatctactggattcaaaaacaattcaagggagctgggctggatctcttatccatca 60
         ||||||
Db   253 AATGGG 258

Qy   61 catggg 66
         ||||||
Db   169 GTGAACTTGTTGGATACATCACCATCCATTGAGAGACTCGGGGCTGGCTCACGTATCCGGCT 228

RESULT 12
LOCUS   MMU72207   4737 bp   mRNA   ROD   23-DEC-1996
DEFINITION Mus musculus Eph-and Elk-related kinase (eek) mRNA, complete cds.
ACCESSION   U72207
VERSION   U72207.1 GI:1750258
KEYWORDS   .
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 4737)
  AUTHORS   Park,S.
  TITLE   The Eek receptor, a member of the Eph family of tyrosine protein
            kinases, can be activated by three different Eph family ligands
  JOURNAL   Oncogene (1996) In press
REFERENCE   2 (bases 1 to 4737)
  AUTHORS   Park,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-SEP-1996) Molecular Oncology, Bristol-Myers Squibb,
            P.O. Box 4000, Princeton, NJ 08543, USA
FEATURES   Location/Qualifiers
     source   1..4737
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene   79..3093
                     /gene="eek"
     CDS   79..3093
                     /gene="eek"
                     /note="Eek"
                     /codon_start=1
                     /product="Eph-and Elk-related kinase"
                     /protein_id="AAB39218.1"
                     /db_xref="GI:1750259"
                     /translation="MAPARRLSPALWVTAAAAATCVSAGRGEVNLLDTSTIHGDWG
                     WLTYPARGWDSINEVDESFRPIHTYQVCNVMSPNQNNWLRTNWVPRDGARVYAEIKF
                     TLRDCNSIPGVLGTCKETFNLHYLESDRDLGASTOESOFLRIDTIAADESFTGADLGV
                     RRLKINTEVRGVGPLSKRGFYLAFQDIGACLAILSLRIYYKKCPAMVRNLAAFSEAVT
                     GADSSSLVEVRGQCVRHSEERDTPKMYCSAEGEWLVPIGKCVCSAGYEERRDACMACE
                     LGFYKSAPGDQLCARCPPHSHSATPAAQTCRCDLSYYRAALDPPSAACTRPPSAPVNL
                     ISSVNGTSVTLEWAPPLDPGGRSDITYNAVCRRCPWALSHCEACGSGTRFVPQQTSLA
                     QASLIVANLLAHMNYSFWIEAVNGVSNLSPEPSRAAVVNITNQAPSQVVIRQERA
                     GQTSVSLLWQEPEQPNGIILEYEIKYKEKDKEMQSYSTLRAVTNTRATVSGLKPGTRYV
                     FQVRARTSAGGGRFSQAMEVETGKPRPRYDTRTIWWICLLITIGLVLLLLLICKRRH
                     CGYSKAFQDSDEEKMHYQNGQAPPVPVFLPLNHPPGKFPETOFSAEPHTYEEPGRAGRS
                     FTREIEASRIHIEKIIGSGESGEVCYGRLCVPGQGORDVPVAIKALRAGYTERQRODFLS
                     EAAIMGQFDHPNIIRLEGVVTRGRLAMIVTEYMENGSLDAFLRTHDQOFTIVQLVGML
                     RGVGAGMRYLSDLGIYIHRDLAARNVLVDGRLVCKVSDFGLSRALEDDPEAAYTTAGGK
                     IPIRWTAPEAIAFRFTSSASDVWSFGVVMWEVLAYGERPYWNMTNQDVISSVEEGYRL
                     PAPMGCPRALHQLMLDCWHKDRAORPRFAHVSVLDALVHSPESLRATAATVSRCPPPA
                     FARSCPDLRAGSGNGDLTVDWLDSIRMGRYRDHFAAGGYSSLGMVLRMNAQDVRAL
                     GITLMGHQKKILGSIQTMRAQLSSIQGRRHI."
BASE COUNT   971 a   1504 c   1381 g   881 t
ORIGIN

Qy   1 gtcaatctactggattcaaaaacaattcaagggagctgggctggatctcttatccatca 60
         |||||||||||||||||||||||||||| |||||||||||||||||| ||||| |||||
Db   229 CATGGG 234

Query Match   49.1%; Score 32.4; DB 4; Length 4124;
Best Local Similarity 68.2%; Pred. No. 0.033;
Matches   45; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Query Match   49.1%; Score 32.4; DB 12; Length 4737;
Best Local Similarity 68.2%; Pred. No. 0.033;
Matches   45; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
RESULT 13
LOCUS       HS61A9        160703 bp    DNA            PRI       22-NOV-1999
DEFINITION  Human DNA sequence from clone RP1-61A9 on chromosome 1p35.2-36.13,
            complete sequence.
ACCESSION   AL035703
VERSION     AL035703.20  GI:6434643
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160703)
  AUTHORS   Wilson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 15, 1999 this sequence version replaced gi:6114770.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP1-61A9 is from the library RPCI-1 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP1-61A9.
FEATURES             Location/Qualifiers
     source          1..160703
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p35.2-36.13"
                     /clone="RP1-61A9"
                     /clone_lib="RPCI-1"
     misc_feature    complement(939..1245)
                     /note="match: GSS: Em:AQ134928"
     misc_feature    9482..9888
                     /note="match: GSS: Em:AQ010082"
     misc_feature    25683..25889
                     /note="Single clone region"
     misc_feature    29241..29448
                     /note="match: GSS: Em:AQ594605.1"
     misc_feature    29451..29575
                     /note="Weak Data"
     misc_feature    41090..41437
                     /note="match: GSS: Em:AQ816982.1"
     misc_feature    complement(51292)
                     /note="Tandem repeat. Forced join. Gap sized by
                     Restriction digest to be 1200-1500bp"
     misc_feature    complement(82130..82878)
                     /note="match: GSS: Em:AQ750654.1"
     misc_feature    complement(82496..82871)
```

```
     misc_feature    82591..82878
                     /note="match: GSS: Em:AQ816014.1"
     misc_feature    86392..86712
                     /note="match: STS: Em:HSA048YH1"
     misc_feature    118067..118469
                     /note="match: GSS: Em:AQ670525.1"
     misc_feature    complement(121941..122454)
                     /note="match: GSS: Em:AQ210441"
     misc_feature    126813..127149
                     /note="match: STS: Em:HSPE55G06"
     misc_feature    126867..127465
                     /note="match: GSS: Em:AQ260557"
     misc_feature    126892..127225
                     /note="match: GSS: Em:AQ419366"
     misc_feature    130356..130685
                     /note="match: GSS: Em:AQ264734"
     misc_feature    130356..130674
                     /note="match: GSS: Em:AQ766010.1"
     misc_feature    130376..130747
                     /note="match: GSS: Em:AQ332261 Em:AQ332293"
     misc_feature    130586..131194
                     /note="match: GSS: Em:AQ530794.1"
     misc_feature    complement(138395..138703)
                     /note="match: GSS: Em:AQ746639.1"
     misc_feature    complement(138397..138766)
                     /note="match: GSS: Em:AQ214151"
     misc_feature    complement(138397..138744)
                     /note="match: GSS: Em:AQ429746.1"
     misc_feature    complement(138398..138763)
                     /note="match: GSS: Em:AQ504997.1"
     misc_feature    138773..139200
                     /note="match: GSS: Em:AQ718784.1"
     misc_feature    138774..139225
                     /note="match: GSS: Em:AQ118462"
     misc_feature    139230..139556
                     /note="match: GSS: Em:AQ700467.1"
     misc_feature    139243..139556
                     /note="match: GSS: Em:AQ876817.1"
     misc_feature    147098..147635
                     /note="match: GSS: Em:AQ602397.1"
     misc_feature    complement(152202..152620)
                     /note="match: GSS: Em:AQ479682.1"
     misc_feature    complement(152620)
                     /note="match: STS: Em:G58359.1; match: GSS: Em:AQ347856.1"
BASE COUNT    41742 a  40589 c  39795 g  38577 t
ORIGIN
```

```
Query Match    49.1%;  Score 32.4;  DB 10;  Length 160703;
Best Local Similarity   68.2%;
Matches   45;  Conservative  0;  Mismatches  21;  Indels  0;  Gaps  0;

Qy          1 gtcaaatctaactggaattcaaaacaattcaaaggggagctgggctggagtcctcttatccatca 60
              || |||||| | | ||||| | | ||| | |||||||| | ||||| |||| || ||| | ||||
Db      59996 GTGAATTTGCTGGACACGTCGACCATCCACGGGGACTGGGCTGGGCTCACGTATCCGGCT 59937

Qy         61 catggg 66
              ||||||
Db      59936 CATGGG 59931

RESULT 14
LOCUS       HUMRPTK       4523 bp    mRNA            PRI       10-AUG-1995
DEFINITION  Homo sapiens receptor protein-tyrosine kinase (HEK11) mRNA,
            complete cds.
ACCESSION   L36642
VERSION     L36642.1  GI:551607
KEYWORDS    EPH-like receptor PTK; receptor protein-tyrosine kinase.
SOURCE      Homo sapiens female fetus, 17-18 weeks gestation brain cDNA to
            mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE   1 (bases 1 to 4523)
  AUTHORS   Fox,G.M., Holst,P.L., Chute,H.T., Lindberg,R.A., Janssen,A.M.,
             Basu,R. and Welcher,A.A.
  TITLE   cDNA cloning and tissue distribution of five human EPH-like
             receptor protein-tyrosine kinases
  JOURNAL   Oncogene 10 (5), 897-905 (1995)
  MEDLINE   95206782
FEATURES   Location/Qualifiers
     source   1..4523
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="fetus, 17-18 weeks gestation"
                     /sex="female"
                     /clone="Stratagene premade library, cat #936206"
                     /tissue_type="brain"
     mRNA   1..4523
                     /gene="HEK11"
                     /product="receptor protein-tyrosine kinase"
     gene   1..4523
                     /gene="HEK11"
     5'UTR   1..185
                     /gene="HEK11"
     CDS   186..3182
                     /gene="HEK11"
                     /codon_start=1
                     /product="receptor protein-tyrosine kinase"
                     /protein_id="AAA74243.1"
                     /db_xref="GI:551608"
                     /translation="MVFQTRYPSWIILCYIWLLRFAHTGEAQAAKEVLLLDSKAQQTE
                     LEWISSPPNGWEEISGLDENYTPIRYRYQVCQVMEPNQNNWLRTNWISKGNAQRIFVEL
                     KFTLRDCNSLPGVLGTCKETENLYYYETDTDIGRNIRENLYVKIDTIAADESFTQGDL
                     GERKMKLNTEVREIGPLSKKGFYLARQDVGACIALVSVKYYKKCWSIIENLAIFPDT
                     VTGSEFSSLVEVRGTCVSSAEEEAENAPRMHCSAEGEWLVPIGKCICKAGTQQKGDTC
                     EPCGRGFYKSSSSQDLQCSRCPTHSFSDKEGSSRCECEDGYYRAPSDPPYVACTRPPSA
                     PQNLIFNINQTTVSLEWSPPADNGGRNDVTYRILCKRCSWEQGECVPCGSNIGYMPQQ
                     TGLEDNYVTVMDLLAHANYTFEVEAVNGVSDLSRSQRLFAAVSITTGQAAPSQVSGVM
                     KERVLQRSVELSWQEPEHPNGVITEYEIKYYEKDQRERTYSTVKTKSTSASINNLKPG
                     TVYVFQIRAFTAAGYGNYSPRLDVATLEEATGKMEEATAVSSEQNPVIIAVVAVAGT
                     IILVFMVFGFIIGRRHCGYSKADQEGDEELYFHFKFPGTKTYIDPETYEDPNRAVHQF
                     AKELDASCIKIERVIGAGEFGEVCSGRLKLPGKRDVAVAIKTLKVGYTEKORDFLCE
                     ASIMGQFDHPNVVHLEGVVTRGKPVMIVIEFMENGALDAFLRKHDGQFTVLQLVGMLR
                     GIAAGMRYLADMGYVHRDLAARNILVNSNLVCKVSDFGLSRVIEDDPEAVYTTGKI
                     PVRWTAPEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMSNQDVIKAIEEGYRLP
                     APMDCPAGLHQLMLDCWQKERAERPKFEQIVGILDKMIRNPNSLKITPLGTCSRPISPL
                     LDQNTPDFTIPCSVGEWLQAIKMERYKDNFTPAAGYNSLESVARMTIEDVMSLGITLVG
                     HQKKIMSSIQTMRAQMLHLHGTGIQV"
     3'UTR   3183..4523
                     /gene="HEK11"
BASE COUNT   1447 a   833 c   999 g   1244 t
ORIGIN

Query Match   47.38; Score 31.2; DB 10; Length 4523;
Best Local Similarity   70.0%; Pred. No. 0.097;
Matches   42; Conservative   0; Mismatches   18; Indels   0; Gaps   0;

Qy   7 ctactggattcaaaaacaattcaaggggagctggacctggatcctcttatccatcacatggg 66
         |||||||||||||||||||| |||| |||| ||| ||||| ||||| |||||||||||||
Db   288 CTGCTGGATTCTRAAAGCACAACAACACAGAGTTGGAGTGGAGGATTCCTCTCCACCAATGGG 347

RESULT 15
HSJ189K14
LOCUS   HSJ189K14 116490 bp   DNA   HTG   23-NOV-1999
DEFINITION Homo sapiens chromosome 6 clone RP1-189K14, *** SEQUENCING IN
            PROGRESS ***, 6 unordered pieces.
ACCESSION   AL121966
VERSION   AL121966.1 GI:6981849
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 116490)
  AUTHORS   Sims,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-OCT-1999) Wellcome Trust Genome Campus, Hinxton,
             Cambridgeshire, CB10 1SA, UK. E-mail enquiries: cloherequest@sanger.ac.uk
             humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  COMMENT   On Feb 16, 2000 this sequence version replaced gi:6066105.
             IMPORTANT: This sequence is unfinished and does not necessarily
             represent the correct sequence. Work on the sequence is in progress
             and the release of this data is based on the understanding that the
             sequence may change as work continues. The sequence may be
             contaminated with foreign sequence from E.coli, yeast, vector,
             phage etc. Order of segments is not known; 800 n's separate
             segments. Unfinished: dJ189K14 Contig_ID: 00219 acc=AL121966
             Length: 4081 bp Unfinished: dJ189K14 Contig_ID: 00235
             acc=AL121966 Length: 19639 bp Unfinished: dJ189K14 Contig_ID:
             00870 acc=AL121966 Length: 26994 bp Unfinished: dJ189K14
             Contig_ID: 00919 acc=AL121966 Length: 31021 bp Unfinished:
             dJ189K14 Contig_ID: 01165 acc=AL121966 Length: 29139 bp
             Unfinished: dJ189K14 Contig_ID: 01338 acc=AL121966 Length: 1616
             bp.
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 6 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *   1   4081: contig of 4081 bp in length
             *   4082   4881: gap of   800 bp
             *   4882   24520: contig of 19639 bp in length
             *   24521   25320: gap of   800 bp
             *   25321   52314: contig of 26994 bp in length
             *   52315   53114: gap of   800 bp
             *   53115   84135: contig of 31021 bp in length
             *   84136   84935: gap of   800 bp
             *   84936   114074: contig of 29139 bp in length
             *   114075   114874: gap of   800 bp
             *   114875   116490: contig of 1616 bp in length.
FEATURES   Location/Qualifiers
     source   1..116490
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP1-189K14"
                     /clone_lib="RPCI-1"
BASE COUNT   37770 a   18571 c   20011 g   36038 t   4000 others
ORIGIN

Query Match   47.38; Score 31.2; DB 32; Length 116490;
Best Local Similarity   70.0%; Pred. No. 0.13;
Matches   42; Conservative   0; Mismatches   18; Indels   0; Gaps   0;

Qy   7 ctactggattcaaaaacaattcaaggggagctggacctggatcctcttatccatcacatggg 66
         |||||||||||||||||||| |||| |||| ||| ||||| ||||| |||||||||||||
Db 15905 CTGCTGGATTCTRAAAGCACAACAACACAGAGTTGGAGTGGAGGATTCCTCTCCACCAATGGG 15964

Search completed: May 15, 2000, 11:43:14
Job time: 18621 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 11:20:19 ; Search time 1358.94 Seconds
                (without alignments)
                1968.540 Million cell updates/sec

Title: US-09-104-340-8
Perfect score: 660
Sequence: 1 tgggaagagatcaagtggtgt..........aaagaggtttatgtgccaa 660

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   EST:*

1: em_est1:*	45: gb_est26:*	73: gb_est43:*	99: em_gss9:*
2: em_est2:*	46: gb_est27:*	74: gb_est44:*	100: em_gss10:*
3: em_est3:*	47: gb_est28:*	75: em_est31:*	101: em_gss11:*
4: em_est4:*	48: gb_est29:*	76: gb_gss4:*	102: gb_gss10:*
5: em_est5:*	49: gb_est30:*	77: em_est32:*	103: gb_gss11:*
6: em_est6:*	50: gb_est31:*	78: em_est33:*	104: em_gss12:*
7: em_est7:*	51: gb_est32:*	79: em_est34:*	105: gb_gss12:*
8: em_est8:*	52: gb_est33:*	80: gb_est45:*	106: gb_gss13:*
9: em_est9:*	53: em_est21:*	81: gb_est46:*	107: gb_gss14:*
10: em_est10:*	54: em_est22:*	82: gb_est47:*	108: gb_gss15:*
11: em_est11:*	55: gb_est34:*	83: gb_gss1:*	109: gb_gss16:*
12: em_est12:*	56: gb_est35:*	84: gb_gss2:*	
13: em_est13:*	57: gb_est36:*	85: gb_gss3:*	
14: em_est14:*	58: em_est26:*	86: gb_gss4:*	
15: em_est15:*	59: gb_est37:*	87: em_gss1:*	
16: em_est16:*	60: gb_est38:*	88: em_gss2:*	
17: em_est17:*	61: gb_est35:*	89: em_gss3:*	
18: em_est18:*	62: gb_est36:*	90: em_gss4:*	
19: em_est19:*	63: gb_est37:*	91: gb_gss5:*	
20: gb_est1:*	64: gb_est38:*	92: gb_gss6:*	
21: gb_est2:*	65: em_est27:*	93: gb_gss7:*	
22: gb_est3:*	66: em_est28:*	94: gb_gss8:*	
23: gb_est4:*	67: em_est29:*	95: em_gss5:*	
24: gb_est5:*	68: em_est30:*	96: em_gss6:*	
25: gb_est6:*	69: gb_est39:*	97: em_gss7:*	
26: gb_est7:*	70: gb_est40:*	98: em_gss8:*	
27: gb_est8:*	71: gb_est41:*		
28: gb_est9:*	72: gb_est42:*		
29: gb_est10:*			
30: gb_est11:*			
31: gb_est12:*			
32: gb_est13:*			
33: gb_est14:*			
34: gb_est15:*			
35: gb_est16:*			
36: gb_est17:*			
37: gb_est18:*			
38: gb_est19:*			
39: gb_est20:*			
40: gb_est21:*			
41: gb_est22:*			
42: gb_est23:*			
43: gb_est24:*			
44: gb_est25:*			

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                   %   SUMMARIES
Result   Query

No.	Score		Match Length	DB ID		Description
1	276.6	41.9	835	69	AU080921	AU080921
2	200.6	30.4	435	23	H38363	H38363 YP50g04.r1
3	184.2	27.9	572	74	AW175051	AW175051 f131g02.y
4	143.4	21.7	424	84	B36983	B36983 HS-1042-a2-
5	130.8	19.8	474	41	AI011631	AI011631 EST206082
6	129.8	19.7	473	41	AI011632	AI011632 EST206083
7	119.2	18.1	406	21	R15219	R15219 yf88e12.r1
c 8	113.4	17.2	408	44	AI325344	AI325344 m101c02.y
c 9	110.2	16.7	366	27	AA024123	AA024123 m101c02.r
10	97.3	14.7	493	44	AI249967	AI249967 qx47d02.x
11	93	14.1	348	44	AI323334	AI323334 m101c02.x
12	92.6	14.0	496	63	AI337308	AI337308 tb98b01.x
13	88.2	13.4	344	80	AI967980	AI967980 wu1lf04.x
14	86	13.0	344	64	AW045268	AW045268 UI-M-BH1-
15	75.4	11.5	314	64	AW366937	AW366937 UI-M-BH1-
16	75.4	11.4	296	23	H25851	H25851 yl53g03.s1
17	70.2	10.6	400	79	AW233440	AW233440 fj32e03.s1
c 18	68.6	10.4	800	61	AL040428	AL040428 DKFZp434C
c 19	68.6	10.5	687	61	AL040428	AL040428 DKFZp434C
c 20	62.8	9.4	375	62	T78739	T78739 yd0le03.r1
21	58.4	8.8	572	62	AI878971	AI878971 au51h08.y
c 22	55.6	8.4	221	61	AL079781	AL079781 DKFZp434M
c 23	54.4	8.2	628	41	AI003549	AI003549 a188a09.s
24	54	8.2	151	60	AI794583	AI794583 fc46b08.y
c 25	50.8	7.7	416	61	AF071057	AF071057 DKFZp434N
c 26	50.8	7.7	474	61	AL045967	AL045967 DKFZp434N
27	48.8	7.4	643	103	AQ201571	AQ201571 RPCII1-58
28	40.2	6.1	1099	83	AL104186	AL104186 Drosophil
c 29	39.8	6.0	494	59	AI767989	AI767989 wj01d11.x
c 30	39.8	6.0	1063	83	CNS01517	CNS01517 Drosophil
31	39	5.9	552	61	AI819882	AI819882 w105b07.x
32	37.4	5.7	324	20	M79194	M79194 EST01342 Su
c 33	37.2	5.7	802	105	AQ330023	AQ330023 nbxb0046K
c 34	35.4	5.5	566	108	AQ526929	AQ526929 CITBI-E1-
c 35	35.4	5.4	237	84	B47312	B47312 HS-1056-A1-
36	35	5.4	1101	83	CNS0182P	CNS0182P Drosophil
37	35	5.3	462	63	AW017395	AW017395 EST272817
38	34.8	5.3	637	79	AW256663	AW256663 EST304800
39	34.6	5.2	459	45	AI366919	AI366919 qv94h06.x
c 40	34.6	5.2	629	42	AI143874	AI143874 qe04f05.x
c 41	34	5.2	320	42	AI112217	AI112217 CNS010x6x
c 42	34	5.2	1101	82	AI098595	AI098595 Drosophil
c 43	33.8	5.1	246	94	AQ036415	AQ036415 CIT-HSP-2
44	33.8	5.1	409	48	AA967414	AA967414 ua05e02.r
45	33.8	5.1	458	48	AI551450	AI551450 mp86g03.y

```
RESULT   1
LOCUS       AU080921        835 bp    mRNA            EST       20-OCT-1999
DEFINITION  AU080921 Sugano mouse brain mncb Mus musculus cDNA clone MNCb-6241
            5', mRNA sequence.
ACCESSION   AU080921
VERSION     AU080921.1  GI:6085675
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 835)
AUTHORS     Hashimoto,K., Kusuda,J., Toyoda,A., Tanuma,R., Ito,A., Hirata,M.,
            Suzuki,Y., Sasaki,M. and Sugano,S.
TITLE       Isolation of full-length cDNA clones from a mouse brain cDNA
            library made by oligo-capping method
JOURNAL     Unpublished (1999)
COMMENT     On May 18, 1998 this sequence version replaced gi:3137227.
            Contact: Katsuyuki Hashimoto
            Division of Genetic Resources
            National Institute of Infectious Diseases
            23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            Email: khashi1@nih.go.jp
            URL: http://www.nih.go.jp/yoken/genbank/.
FEATURES             Location/Qualifiers
     source          1..835
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="MNCb-6241"
                     /clone_lib="Sugano mouse brain mncb"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="TOP10"
                     /note="Organ: brain; Vector: pME18S-FL3; 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [AIGTGGCCTTTTTTTTTTTTTTTTT] double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGGCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3.
                     XhoI sites just outside the DraIII sites can be used to
                     isolate the cDNA insert. Size selection was performed to
                     exclude the fragments <1.5 kb. Library was constructed by
                     Sugano et al. (University of Tokyo, Institute of Medical
                     Science). Custom primer for sequencing: 5' end primer
                     [CTTCTGCTCTAAAAGCTGCG]."
BASE COUNT      202 a    182 c    230 g    190 t    31 others
ORIGIN
```

ALIGNMENTS

```
Query Match   41.9%;  Score 276.6;  DB 69;  Length 835;
Best Local Similarity   65.9%;  Pred. No. 6.5e-71;
Matches 396;  Conservative   0;  Mismatches 200;  Indels   5;  Gaps   1;

Qy     1 tgggaaggatcagtgtgtggatgaacattacacaccccatcaggagactacaggtgtgc 60
         ||||| |||  | || | || || | || ||||| |   |||||| || |||||||| ||
Db   216 TGGGAGGAAGTGAGCATTATGGAATAGAGAAAATTACACCGGTTCCGAACCTACCAGGTGTGC 275

Qy    61 aatgtcatggaccacagtcaaacaaactggttgctgaagaacaaacctgggttcccaggaactaa 120
         |  ||| ||||| |||| | |||||| || |||  ||||||  | ||| ||  ||| ||| || 
Db   276 AACGTGATGGAAGCCAGCCAGAACAACTGGCTCGGAACTGACTGGGATCACCCGCGAAGGG 335

Qy   121 gctcagaagagattatgtggagctcaagttcactcctacgagactgcatagcattccatty 180
         || ||| ||| |||| ||  |  |||||||||| ||||  | |||| ||| ||| |  || 
Db   336 GCGCAGAGGGTGTCATTGAAATTAAGTTCACTCTAAGGACTGCCAACAGTCTTCCGGGC 395

Qy   181 gtttaggaacttgcaaggagacattcaacctgtactacatgagtctgatgatgatcat 240
         ||   |||||||||||||||||| || ||| ||||||| |||||| || |||| || 
Db   396 GTCATGGGGGACTTGCAAGGAGACGTTTAACCTCTACTACTATGAGTCGGACAACGACAAG 455

Qy   241 ggggtgaaatttcgagaagatcagtttacaaaagattgcaccattgcagcctgatgaaagt 300
         |   |  ||| |||||| || |||| ||| ||| |  |||||| ||| | |||| |||| 
Db   456 GAGGGCTTCATCAGAGAGAGCAGTTTGGCAAGATCGACACACATCGGAGCCGAGCTGAAGAGC 515

Qy   301 ttcactcaaatggatctgggggaccgtattctgaagctcaaacaacactgagattagaagagta 360
         ||||| ||| ||||| ||| | ||||  |||| | |||||| ||||||||| || | | |||| 
Db   516 TTCACGCAGGTGGACATTGGTGACGGAATCATGAAACTCAACACTGAGATCCGGGATGTA 575

Qy   361 ggtccctgtcaacaagaaggagattcttattttggacattcaagatgttggttcgtgttcc 420
         || || || ||||||| ||  ||| | | ||  ||| ||||||||||| ||| |||| ||||
Db   576 GGGCCACTGAGCAACGAAGAAAGGGGTTTTAACCTTGGCGTTTTCAGGATGTGGGGGTTGGC 635

Qy   421 ttggtgtctgtggagtatacctcaaaaagtgcccaatttacagtgaagaaatctggctatg 480
         |||||||||  | | | |||  ||||| || || ||| | || |  |  || | || || 
Db   636 CTGGTGTCTTCGTCCNNTGGTTCTACAANAAGTGNCCACTTNCCCGTTCNAAAATTANCCCAN 695

Qy   481 ttccagacacgtacccatggact-----cccagtccctgtgaggttagaggtctt 535
         ||||  | ||  || ||  || |     ||||||| | |  |  ||||||| | |
Db   696 TTTCCTGATACCATTACTCGGGCGGTTCATACCCTCTTTNCCTGGNGAAGGTTCNAGGCTCCT 755

Qy   536 gtgtcaaacaattctaaggaggaggatcctccaggagtactccagtactaaggagagaggcgaat 595
         | ||  ||| | | | ||||||||||| | || |||  |||| ||||| | |||| | || |  
Db   756 GNGTNAACNACTNANANAAAAAAAAAAGGATGNGCCNAAAATGNNCTGNGGGGGCNNANGGNNANG 815
```

```
Qy    596 g 596
Db    816 G 816
```

RESULT 2
H38363
H38363
```
LOCUS       H38363      435 bp    mRNA            EST      16-AUG-1995
DEFINITION  yp50g04.r1 Soares retina N2b4HR Homo sapiens cDNA clone
            IMAGE:190902 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
            RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   H38363
VERSION     H38363.1   GI:907862
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 435)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     On May 5, 1995 this sequence version replaced gi:797807.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 276
            Insert Size: 2663
            High quality sequence stops: 276
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2663     Std Error: 0.00
            Seq primer: M13Rev
            High quality sequence stop: 276.
FEATURES             Location/Qualifiers
     source          1..435
                     /organism="Homo sapiens"
                     /db_xref="GDB:3847311"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:190902"
                     /clone_lib="Soares retina N2b4HR"
                     /sex="male"
                     /tissue_type="retina"
                     /dev_stage="55 year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAACTGGAAGCGGCGGCCGCCGCGTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). The retinas were obtained from a 55 year old
                     caucasian and total cellular poly(A)+ RNA was extracted 6
                     hrs after their removal. The retina RNA was kindly
                     provided by Roderick R. McInnes M.D. Ph.D. from the
                     University of Toronto. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      109 a    108 c    103 g    109 t     6 others
ORIGIN
```

```
Qy    1    tgggaagagatcagtgtgtgatgaacattacacacccataggagactaccaggtgtgc        60
           ||||||||||||||||| ||||| ||||||||||||||||||||||| |||||||||||
Db    220  TGGGAAGAGATCAGTGGTGTGGATGAACATTACACACCCATAGGAGACTTACCAGGTGTGC     279

Qy    61   aatgtcatggaccacagtcaaaacaattggcttgagagaacaaactgggtccccaggaactca    120
           ||| ||| ||||||||||||| |||||| ||||| ||||||||||||||||||||||||||||
Db    280  AAYGTCAYGGACCACAGTCAYYAACAATTGGCTGAGAGAACAAACTGGGTCCCCAGGAACTCA     339

Qy    121  gcttcagaagaattatgtggagctcaagttcacttctacgagactgcaatagcattccattg    180
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db    340  GCTTCAGAAGAATTTATGTGGAGCTCAAGTTCACTTCTACGAGACTGCAATAGCCATTCCATTG    399

Qy    181  gtttttaggaaacttgcaaggagagacattcaac     210
           |||||||||||  |||||||  |||||| ||||||
Db    400  GTTTTTAGGNACTTNCAAGGGGACATTAACC          429
```

RESULT 3
AW175051
AW175051
```
LOCUS       AW175051       572 bp    mRNA            EST      16-NOV-1999
DEFINITION  f131g02.y1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
            2639282 5' similar to SW:EPA4_CHICK Q07496 EPHRIN TYPE-A RECEPTOR 4
            PRECURSOR ;, mRNA sequence.
ACCESSION   AW175051
VERSION     AW175051.1   GI:6441079
KEYWORDS    EST.
SOURCE      zebrafish.
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
            Cyprinoidea; Cyprinidae; Rasborinae; Danio.
REFERENCE   1  (bases 1 to 572)
AUTHORS     Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B.,
            Ritter,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE       WashU Zebrafish EST Project 1999
JOURNAL     Unpublished (1999)
COMMENT     On Jun 5, 1998 this sequence version replaced gi:3189572.
            Other_ESTs: f131g02.x1
            Contact: S.L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 485.
FEATURES             Location/Qualifiers
     source          1..572
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone="2639282"
                     /clone_lib="Sugano Kawakami zebrafish DRA"
                     /sex="mixed (one male and one female, including
                     unfertilized eggs)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer [ATGTGGCCGCTTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.    Size selection was performed to exclude fragments
                     <1.5kb.    Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science) and
                     kindly donated by Dr. Koichi Kawakami. Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     sequencing:
```

RESULT 4

```
BASE COUNT      169 a     136 c     160 g     104 t
ORIGIN

Query Match      primer CGACCTGCAGCTCGAGCACA.               "      3 others

Best Local Similarity    70.8%;  Score 184.2;  DB 74;  Length 572;
Matches  257;  Conservative   0;  Mismatches 105;  Indels   1;  Gaps  1;

Qy       1 tgggaagagatcagtggtggatgaacattacacaccatca-ggacttacaggtgtg         59
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     196 TGGGAAGAGATCAGTGGTGGATGAACATTACACACCATCAGGGACTTACAGGTGTG        255

Qy      60 caatgtcatggcacacagtcaaaaacaattggctggagagacaaactgggtccccaggaactc       119
           |||||| ||||| ||| |||||||||||| |||||||||||||||||||||||||||||||||
Db     256 CAACGTAATGGCACCAGCCACAACCAGCTCCGGACACGGAGATCCCCGTGGGCG        315

Qy     120 agctcagaagagattatgtggagctcaagttcactctacgagagactgcataagcattccatt       179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 AGCTCAGCGGGGTCACGAGATCAAATTCACCCTGCCTGACTGCAACAGCCTGCCAGG       375

Qy     180 ggttttaggaacttgcaaggagaacaattcaaacctgtacttacacatggagtctgatgatcca       239
           |||||||||||  |||||||||||||||||||||||||||||||||| |||  ||||||||
Db     376 AGTTCATGGGACCTGCAAGGAGACCTTAATCTTTACTATTATGAGTCCAACAATGACAA       435

Qy     240 tgggttgaaattccgagagcatcagtttacacaagattgacacccattgcaagctggatgaaag       299
           ||||||||||||||||||  ||||||| ||||||||||||||| |||||||||
Db     436 AGACGATACATCCGAGAAAAACCAGTTCACCAAAAATGACACACATTGCAGCTGATGAAAG        495

Qy     300 ttccacttcaaatggatcttgggacccgtattctgaagctcaacactgagagttgagaagt       359
           ||  ||| ||||||||| |||||| |||||||||||||||||||||||
Db     496 CTTCACGCAGGTGGACATAGGAGACACCGCATCATCATGAAGCTGAAACACNGAGGTGAGGGGATGT       555

Qy     360 agg 362
           |||
Db     556 GGG 558
```

```
LOCUS       B36983        424 bp     DNA                GSS      17-OCT-1997
DEFINITION  HS-1042-A2-F05-MR.ab1 CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 824 Col=10 Row=K, genomic survey
            sequence.
ACCESSION   B36983
VERSION     B36983.1    GI:2536352
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Trainoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 824  row: K  column: 10
            Class: BAC ends
            High quality sequence stop: 424.
FEATURES             Location/Qualifiers
     source          1..424
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 824 Col=10 Row=K"
```

RESULT 5

```
BASE COUNT      123 a     77 c     101 g     122 t                1 others
ORIGIN

Query Match

Best Local Similarity    67.7%;  Score 143.4;  DB 84;  Length 424;
Matches  233;  Conservative   0;  Mismatches 102;  Indels   9;  Gaps  2;

Qy     326 gtattctgaagctcaacactgagattaagagaagtaggtcctgtcaacagagaaggggattt       385
           || |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       2 GNATACTCAAGCTTAACACTGAGGTTGGACCTTTGTCCAAAAGAAAGGGGGATTCT        61

Qy     386 atttggcattcaagatgttggtcctggtgttgttggctgttggtgagagtactctca       445
           ||||||||||||||||| ||| |||||||||||||||||||||||||||
Db      62 ATCTTGCCTTTCAGGATGTRAGGGGCTTGCATAGCTTTGGTTTCTCGTCAAAGTGTACTACA       121

Qy     446 aaaagtgcccattacagtgaagaatctggctatgtttccagacaacggtacc---catgg       502
           |||||||||||||||||||||||| |||||| |||||||||||||||||
Db     122 AGAAGTGCTGGTCCATTATTGAGACTTAGCTATGTTTCCAGATACAGTGACTGACTGGTTCAG       181

Qy     503 actccagtccctgtgtggaggttagaggtttgtgtcaacaattctaagga------gg       556
           ||| |||||||||||||||||||||| |||||||||||||||||||  ||
Db     182 AATTTTCCTCTTTAGTCGAGGTTCGAGGGTTCCGAGGGACATGGTGTCACCCAGGAAGAAGCCG       241

Qy     557 aagacctccaaggatgttactcgcagtacagaaggcgaatggcttgtacccattggcaagt       616
           |||||| |||||||||||||| ||||||||||||||||||| |||||
Db     242 AAAAGCGCCCCCAGGAGATGCACTGCAGATGGTTAGTCGCCATTGGAAAAT       301

Qy     617 gttcctgcaatgctggctatgaagagaagaggtttatgtgccaa       660
           |||| ||| ||||||||||||||||||||||||||||||||||
Db     302 GTATCTGCAAAGCAGCCTACCAGCAAAAAGGAGAGACACTTGTGTGAA       345
```

```
LOCUS       AI011631        474 bp     mRNA               EST      31-JAN-1999
DEFINITION  EST206082 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
            ROVAX55 3' end, mRNA sequence.
ACCESSION   AI011631
VERSION     AI011631.1    GI:3225463
KEYWORDS    EST.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
  TITLE     Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
            Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2153373.
            Contact: Lee, NH
            ATCC
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            Seq primer: M13-21.
FEATURES             Location/Qualifiers
     source          1..474
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
                     /clone="ROVAX55"
                     /note="Organ: ovary; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI"
                     /clone_lib="Normalized rat ovary, Bento Soares"
                     /db_xref="ATCC (inhost):2021827"
```

BASE COUNT
ORIGIN

Query Match   117 a   129 c   135 g   93 t
Best Local Similarity 19.8%; Score 130.8; DB 41; Length 474;
Matches 216; Conservative 60.38; Pred. No. 4.7e-28;
            0; Mismatches 142; Indels 0; Gaps 0;

```
Qy   29 attacacaccatcaggacttaccaggttgtgcaatgtcatggaccacagtcaaaacaatt  88
        |||||  ||||| |||||||  ||  ||||| |||||||  | ||  | || || |||| |
Db  115 ATGACATGCCTATCTACATGTAACTCCGTGTGCCAATGTGGTCGTGTGGCGACCAGGACAACT  174

Qy   89 ggctgagaacaaactgggtcccccaggaactcagctcagaagatttatgtggagctcaagt  148
        |||| || ||||||||||| || | ||||| |||| |||  ||| |||| |||||||||||
Db  175 GGCTCCGTACCAACTGGGTGTGTACCGGGAGGAGGCCCAGCCCATCTTTATTGAGCTCAAGT  234

Qy  149 tcacttacgagactgcaatagcattcccattggttttaggaacttgcaaggagagacattca  208
        ||||  | |||||||||||||||  ||| | | | ||||| |||  ||||||||||| |||||
Db  235 TCACCGTGGGAGACTGCAATAGCTTTCCGGGTGCCGACTCCTGCCAAAGAGACCTTCA  294

Qy  209 acctgtactacatggagtctgatgatgatcatgggtgaaatttcgagagagcatcagttta  268
        |||| |||||| |||||||| || | || |||||||||||||||| |||| ||||||||||
Db  295 ACCTCTACTATGCAGAGTCAGACGTGGACTATGGGTGGGGTGAAATTTCCAGAAGCGCCAGTTCA  354

Qy  269 caaagattgacaccattgcagctgatgaaagttcactcaaatggatcttggggaccgta  328
        |||| || |||||||| | | || ||||| || || ||||||||| || ||||||| |
Db  355 CCAAGATGCAGACACCATCGCCCCTGACGAGAGATCACGGTCAGAGTGACTTCGAGGCTCGCA  414

Qy  329 ttctgaagctcaacactgagattagagaagtaggtcctgtcaacaagaaggggattta  386
        | || ||||| || |||| |||  |||| | |||| || ||  ||| |||||| || |
Db  415 ATGTCAAGCTGAATGTGGAGGAGCGCCATGGTGGGCGCCCCTTACCCGGAAGGGCTTCTA  472
```

RESULT 6

LOCUS   AI011632   473 bp   mRNA   EST   31-JAN-1999
DEFINITION EST206083 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
            ROVAX56 3' end, mRNA sequence.
ACCESSION   AI011632
VERSION   AI011632.1 GI:3225464
KEYWORDS   EST.
SOURCE   Rattus sp.
  ORGANISM Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 473)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
  TITLE   Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
            Gene Index
  JOURNAL   Unpublished (1998)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2153374.
            Contact: Lee, NH
            ATCC
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            Seq primer: M13-21.
FEATURES   Location/Qualifiers
     source   1..473
                     /organism="Rattus sp."
                     /db_xref="ATCC (inhost):2021828"
                     /clone_lib="Normalized rat ovary, Bento Soares"
                     /note="Organ: ovary; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI"
                     /clone="ROVAX56"
                     /db_xref="taxon:10118"
BASE COUNT   116 a   127 c   136 g   93 t   1 others
ORIGIN

Query Match   117 a   129 c   135 g   93 t
Best Local Similarity 19.7%; Score 129.8; DB 41; Length 473;
Matches 215; Conservative 60.2%; Pred. No. 9.2e-28;
            0; Mismatches 142; Indels 0; Gaps 0;

```
Qy   29 attacacaccatcaggacttaccaggttgtgcaatgtcatggaccacagtcaaaacaatt  88
        |||||  ||||| |||||||  ||  ||||| |||||||  | ||  | || || |||| |
Db  117 ATGACATGCCTATCTACATGTAACTCCGTGTGCCAATGTGGTCGTGTGGCGACCAGGACAACT  176

Qy   89 ggctgagaacaaactgggtcccccaggaactcagctcagaagatttatgtggagctcaagt  148
        |||| || ||||||||||| || | ||||| |||| |||  ||| |||| |||||||||||
Db  177 GGCTCCGTACCAACTGGGTGTGTACCGGGAGGAGGCCCAGCCCATCTTTATTGAGCTCAAGT  236

Qy  149 tcacttacgagactgcaatagcattcccattggttttaggaacttgcaaggagagacattca  208
        ||||  | |||||||||||||||  ||| | | | ||||| |||  ||||||||||| |||||
Db  237 TCACCGTGGGAGACTGCAATAGCTTTCCGGGTGCCGACTCCTGCCAAAGAGACCTTCA  296

Qy  209 acctgtactacatggagtctgatgatgatcatgggtgaaatttcgagagagcatcagttta  268
        |||| |||||| |||||||| || | || |||||||||||||||| |||| ||||||||||
Db  297 ACCTCTACTATGCAGAGTCAGACGTGGACTATGGGTGGGGTGAAATTTCCAGAAGCGCCAGTTCA  356

Qy  269 caaagattgacaccattgcagctgatgaaagttcactcaaatggatcttggggaccgta  328
        |||| || |||||||| | | || ||||| || || ||||||||| || ||||||| |
Db  357 CCAAGATGCAGACACCATCGCCCCTGACGAGAGATCACGGTCAGAGTGACTTCGAGGCTCGCA  416

Qy  329 ttctgaagctcaacactgagattagagaagtaggtcctgtcaacaagaaggggatttt  385
        | || ||||| || |||| |||  |||| | |||| || ||  ||| |||||| || |
Db  417 ATGTCAAGCTGAATGTGGAGGAGCGCCATGGTGGGCGCCCCTTACCCGGAAGGGCTTCT  473
```

RESULT 7

LOCUS   R15219   402 bp   mRNA   EST   13-APR-1995
DEFINITION yf88e12.r1 Soares infant brain 1N1B Homo sapiens cDNA clone
            IMAGE:29543 5' similar to gb:M83941 TYROSINE-PROTEIN KINASE
            RECEPTOR HEK PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   R15219
VERSION   R15219.1 GI:769492
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 402)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE   The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT   On Sep 21, 1992 this sequence version replaced gi:276397.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Size: 1695 Source: IMAGE Consortium, LLNL
            High quality sequence stops: 337
            Insert Length: 1695 Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 337.
FEATURES   Location/Qualifiers
     source   1..402
                     /organism="Homo sapiens"
                     /db_xref="GDB:401890"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:29543"
                     /clone_lib="Soares infant brain 1N1B"
                     /sex="female"
                     /dev_stage="73 days post natal"

```
BASE COUNT      108 a     75 c    110 g    108 t    1 others
ORIGIN
```

```
Qy  332  tgaagctcaacactgagattagaagaagtaggtccctgcaacaagaagggatttattgg  391
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6  TGAAGCTTAACACTGAGGTGAGAGAGGAGGATTGGACCTTTGTCAAAGAAGGGATTCTATCTTG  65

Qy  392  catttcaagatgttggtcttgttcgccttggttgtgagtagtatactctcaaaagt  451
         |||||| ||||| |||||  ||||||||||||||||||| |||||  |||||||||||
Db   66  CCTTTCAGGATGTAGGGGCTTGCATAGCCTTTGGTTTCTGTCATAAGTGTACTACAACAGAGT  125

Qy  452  gcccatttacacagtgaagaactctggctatgttttccagacacacggtacc---catggactccc  508
         ||||||  |||||||||||||||||||||||||||||||| ||| ||||||      |||||||||
Db  126  GCTGGGTCCCATTATTGAGAAACTTAGCTATNCTTTCCAGNATACACAGTGACTCGGTTCAGAGAATTT  185

Qy  509  agtccctggttgaggttaggagggtcttgtgtcaacaattctaaggaggaag------at  561
         |||||||||||||||||||||||||||||||||||||||||||||||||||        ||
Db  186  CCTCTTTAGTCGAGGTTCGAGGGACATGTGTCAGCAGTGCGCAGAGGAAGAACGGGAAAACG  245

Qy  562  ccttccaaggatgtactgcagtacagaggcgaatgcggttgtacccattggcaa  614
         |||  ||||||| ||||||||  ||||||||||||||||||||| ||||||||||
Db  246  CCCCCAGGGATGCACTGCCAGTCAGNAAGGAGAGAATGGTTAGTGCCCCATTGGGGAA  298
```

```
RESULT  8
AI325344/c
LOCUS       AI325344       406 bp    mRNA               EST      23-DEC-1998
DEFINITION  mI01c02.y1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:459170 5' similar to gb:U07634 Mus musculus
            receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
ACCESSION   AI325344
VERSION     AI325344.1     GI:4059773
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2152238.
            Contact: Marra M/Mouse EST Project
            Washington University School of Medicine
            Washington-HHMI Mouse EST Project
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276058
            This read is a RESEQUENCE of a previously sequenced mouse clone.
            This read has been verified (found to hit its original self in the
```

```
BASE COUNT      79 a     117 c    109 g    100 t    1 others
ORIGIN
```

```
Qy   32  acacacccatcaggacttaccaggttgtgcaaatgtcatggaccacagtcaaaacaaatggc  91
         |||  ||||||| ||||||||||||||||||||||||||||||| |  |||||||||||||||
Db  290  ACATGCCTTATCACATGTAACTCGGTGTCCAACGTGGTATNCCGGCGACCAGGACAAACTGGC  231

Qy   92  tgagaacaaactgggttcccagggaactccagtccagaagagattatgtgggagctcaagttca  151
         ||||| |||| |||||||| |||||||||||||||||||| |||||||||||||||||||||||
Db  230  TCCCGCACCAACTGGGTGTACCGGAGGAGGCCCGGACATCTTATTGAGCTCAAGTTCA  171

Qy  152  ctctacgagactgcaatagcattccattggtgtttagtgaacttgcaaggagacattcaacc  211
         | ||||||| ||||||| || ||||||||||||||||| |||||||||||  |||||||||||
Db  170  CGGTGCGAGACTGTAACAGCTTCCCGGGTGGCGGCCCACCTCATCGCAAAGAGACCTTCAACC  111

Qy  212  tgtactacatggagtctgatgatcatgggtgaaatttcgagagacatcagtcagttcacaa  271
         ||||| ||||||| ||||||| | |||||||||||||||||||| ||||||||||| |||| |
Db  110  TCTACTATGCCAGAGTCACAGTGGACTATGGCACCACAACTTCACGAAAGCGCCCAGTTCACA  51

Qy  272  agattgacacaccattgcagctgatgaaa  298
         |||||||| |||||||||||||||||||||
Db   50  AGATTGACNCACCATCGCCCCTGACGAGA  24
```

```
RESULT  9
AA024123/c
LOCUS       AA024123       366 bp    mRNA               EST      21-JAN-1997
DEFINITION  mI01c02.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:459170 5' similar to gb:U07634 Mus musculus
            receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
ACCESSION   AA024123
VERSION     AA024123.1     GI:1489009
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
```

correct orientation)
Seq primer: -40RP from Gibco.

/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5'
AACTGGAAGAATTCCGCGGCCGCCGGGAATTTTTTTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```
Query Match       18.1%;   Score 119.2;   DB 21;   Length 402;
Best Local Similarity  67.9%;   Pred. No. 1.2e-24;
Matches 199;  Conservative   0;  Mismatches  84;  Indels  10;  Gaps  2;
```

```
Query Match       17.2%;   Score 113.4;   DB 44;   Length 406;
Best Local Similarity  64.0%;   Pred. No. 5.9e-23;
Matches 171;  Conservative   0;  Mismatches  96;  Indels   0;  Gaps  0;
```

COMMENT
On May 8, 1995 this sequence version replaced gi:800409.
Contact: Maira M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:276058
Seq primer: -28M13 rev2 from Amersham.

FEATURES
    source   Location/Qualifiers
         1. .366
         /organism="Mus musculus"
         /strain="C57BL/6J"
         /db_xref="taxon:10090"
         /clone="IMAGE:459170"
         /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
         /sex="unknown"
         /tissue_type="placenta"
         /dev_stage="adult"
         /lab_host="DH10B"
         /note="Organ:placenta; Vector: pT7T3D-Pac (Pharmacia)
         with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
         1st strand cDNA was primed with a Not I - oligo(dT) primer
         [5'
         TGTTACCAATCTGAAGTGGGAGCGCGGCGGGAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTT
         T3']; double-stranded cDNA was ligated to Eco RI adaptors
         (Pharmacia), digested with Not I and cloned into the Not I
         and Eco RI sites of the modified pT7T3 vector. Library
         went through one round of normalization, and was
         constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT   69 a   103 c   102 g   92 t
ORIGIN

Query Match   16.7%;   Score 110.2;   DB 27;   Length 366;
Best Local Similarity   63.3%;   Pred. No. 5.1e-22;
Matches 169;   Conservative 0;   Mismatches 98;   Indels 0;   Gaps 0;

```
Qy   32 acacacccatcaaggacttaccaggtgtgtcaatgtcatggaccacagtcaaaaacaattggc 91
           |||||||||| ||||||||||||  | |||||  ||  ||||||||||||||||||||||||
Db  290 ACATGCCTAACTACATACATGTACTCGGTGTGCAACGTGTATCCGGCGACCAGGACAACTGCG 231

Qy   92 tgagaacaaactgggtccccaggaaactcagctcagaagatttatgtgggagctcaagtca 151
           |  |  ||||||||||||  |||| |  |||| ||  ||| |||| |||||||||||||
Db  230 TCCGCCACCAACTGGGTGTACCGGGAGGAGCCGAGCCATCTTATTGAGCTCAAGTTCA 171

Qy  152 ctctacggagactgcaatagcattccattggttttaggaacttgcaaggaggagacattcaaacc 211
           || ||  ||||||| ||| || ||||  |||  | ||||   | |||| |  |||| ||||||
Db  170 CGGTGCGGAGACTGTAACAGCTTCCCGGGGTGGCGCCGCAGCTCATGCAAAAGACACCTTCAAACC 111

Qy  212 tgtactacatggagtctgatgatgatcatggggtgaaatttcgaggagcattacaagtttacaa 271
           | |||||| | ||  ||| ||||| |||||  ||| ||| |||||||||   |||| |||  |
Db  110 TCTACTACTATGCAGAGGTCAGAGATGGACTATGGCACCAACTTCCAGAAGCGGCCCAGTTCACCA 51

Qy  272 agattgacacccattggacgctgataa 298
           |||||||||| ||  |  |||||||
Db   50 AGATTGACACCATCGCCCTTGACGAGA 24
```

RESULT 10
AI249967
LOCUS   AI249967   493 bp   mRNA   EST   05-NOV-1998
DEFINITION   qx47d02.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2004483 3'
     similar to gb:M59371 TYROSINE-PROTEIN KINASE RECEPTOR ECK PRECURSOR
     (HUMAN);, mRNA sequence.
ACCESSION   AI249967
VERSION   AI249967.1   GI:3846496
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 493)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
     Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
     Tel: (301) 496-1550
     Email: Robert.Strausberg@nih.gov

FEATURES
    source   Location/Qualifiers
         1. .493
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="IMAGE:2004483"
         /clone_lib="NCI_CGAP_Pan1"
         /tissue_type="adenocarcinoma"
         /lab_host="DH10B"
         /note="Organ: pancreas; pancreas;
         Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
         Vector: pCMV-SPORT6; Site_1: SalI;
         Average insert size 1.72 kb. Life Technologies catalog #:
         11548-013"
         High quality sequence stop: 297.

BASE COUNT   107 a   141 c   142 g   101 t
ORIGIN

Query Match   14.7%;   Score 97.2;   DB 44;   Length 493;
Best Local Similarity   62.5%;   Pred. No. 3.7e-18;
Matches 167;   Conservative 0;   Mismatches 99;   Indels 1;   Gaps 1;   2 others

```
Qy   29 attacacaccatcaaggacttaccaggtgtgtcaatgtcatggaccacagtcaaaaacaatt 88
           |  |||  ||||| ||||||||||  |  |||||   ||  |||||||||||||||||||||
Db  123 ATGACATGCCGATCTACATGTACTCCGTGTGCAACCGTGATGTGTCTGGCGACCAGGACAACT 182

Qy   89 ggctgagaacaaactgggtccccaggaaactcagctcagaagatttatgtgggagctcaagt 148
           ||||  |  |  ||||||||||  ||||| |  |||| ||  ||| ||||  ||||||| |||
Db  183 GGCTCCGCCACCAACTGGGTGTACCGGGAGGAGGCGCTGAGCCGTATCTTCATTGAGCTCAAGT 242

Qy  149 tcacttactgcgagactgcaaatagcaattccatatggttttaggaacttgcaaggaggacattca 208
           || ||||| | ||||||| |||  |  || || | |||| | |||||  | |||| |  |||| |||
Db  243 TTACTGTACGTGACTGCAACACGTTCCCTGGTGGCGCCGCAGTCTCCTCGCAAAGGAGACTNTCA 302

Qy  209 acctgtacta-catggagtctgatgatgatcatggggtgaaatttcgaggagcatcagtt 267
           ||||  |||| ||  |  ||| |||||  || || ||  ||| |||||||||  | |||
Db  303 ACCTCTACTATGCCCCAGAGTCGGACTACGGCACCAACTTTCAGAAGCGGCTGTGTTC 362

Qy  268 acaaagattgacacccattgcagctgat 294
           || |||||||||| |||| ||  | |||
Db  363 ACCAAGATTGACACCCATTGCGCCCCGAT 389
```

RESULT 11
AI323334
LOCUS   AI323334   348 bp   mRNA   EST   23-DEC-1998
DEFINITION   mi01c02.x1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
     clone IMAGE:459170 3' similar to gb:M59371 TYROSINE-PROTEIN KINASE
     RECEPTOR ECK PRECURSOR (HUMAN); gb:U07634 Mus musculus
     receptor-protein tyrosine kinase (MOUSE);, mRNA sequence.
ACCESSION   AI323334
VERSION   AI323334.1   GI:4057763
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1   (bases 1 to 348)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
     Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
     Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
     Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

```
RESULT 12
LOCUS       AI337308      496 bp    mRNA              EST       18-MAR-1999
DEFINITION  tb98b01.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:2062345 3'
            similar to SW:EPB2_HUMAN P29323 EPHRIN TYPE-B RECEPTOR 2 PRECURSOR
ACCESSION   AI337308
VERSION     AI337308.1  GI:4074235
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI_CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            ;contains element MER22 repetitive element ;, mRNA sequence.
FEATURES             Location/Qualifiers
     source          1..496
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2062345"
                     /clone_lib="NCI_CGAP_Co16"
                     /tissue_type="colon tumor, RER+"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Seq primer: ~400P from Gibco
                     Insert Length: 756   Std Error: 0.00
                     High quality sequence stop: 479.
                     cDNA Library Preparation: M. Bento Soares, Ph.D.
                     cDNA Library Arrayed by: Greg Lennon, Ph.D.
                     DNA Sequencing by: Washington University Genome Sequencing
                     Center
                     Clone distribution: NCI_CGAP clone distribution
                     Information can be
                     found through the I.M.A.G.E. Consortium/LLNL at:
                     this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1057416-1061255, and 1144584-1145351).
                     Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT    104 a    141 c    148 g    103 t
ORIGIN

Query Match   14.0%;  Score 92.6;  DB 45;  Length 496;
Best Local Similarity  55.9%;  Pred. No. 8.3e-17;
Matches 204;  Conservative  0;  Mismatches 149;  Indels 12;  Gaps  1;

Qy    76  agtcaaaacaattggctgagaacaaactgggtccccaggggaactcagctcagaagaagattttat  135
          |||||||| ||| |||| ||| ||||| ||| |||  || ||  | |  |||| |||  |||
Db    11  AGCCAGAACAACTGGCTACGGACCAAGTTTATCCGGCGCCCACCGCCATCCAC   70

Qy   136  gtggagctcaagttcactctacgagagactgcaatagcattccattggtgttttaggaacttgc  195
          ||||||  |||||||  | | | | |  | |||| ||| |||||||||||  |||||| ||||||
Db    71  GTGGAGATGAAGTTTTCGGTGCGTGCGTGACTCAGCACGAGCATCCCCAGCCGTGCCTGGCTCCCTGC  130

Qy   196  aaggagacattcaaacctgtactacatggagtctgatgatgatcatg------------gg  243
          |||||||| ||||| |||||||||||||||  | |||||||||||||            ||
Db   131  AAGGAGACCTTCAACCCTCTATTACTATGGAGGGTGACTTGACTTTGACTCGGCGCCACCAAGACCTTC  190

Qy   244  gtgaaatttcggagagcatcagtttacacaagattgacacactgcagctgatgaaagttc  303
          ||||| ||||||||| | |||| |||||  ||| |||||||| || |  |||| ||||||
Db   191  CCAACTGGATGGAGAATCCATGGGTGAAGGTGGATAACCATTGCACCGACCGAGAGCTC  250

Qy   304  acctcaaatggactttgggaccgtatctctgaagctcaacactgagattagagaagtaggt  363
          ||| ||||||||| | ||||| ||||||||||||||||||||||||||| |||||| |||
Db   251  TCCCAGGTGGAGCACCTGGGTGGCGCGGTCATGAAAATCAACACCGAGGTGGAGCTTCGGA  310
```

```
TITLE       Waterston,R.
JOURNAL     The WashU-HHMI Mouse EST Project
            Unpublished (1996)
COMMENT     On Jan 17, 1998 this sequence version replaced gi:2043939.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mousees@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:276058
            This clone was previously sequenced on the 3' end
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Trace considered overall poor quality
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..348
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:459170"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAAATCTGAAGTGGGAGCGGCGGCGGGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT    87 a    94 c    93 g    74 t
ORIGIN

Query Match   14.1%;  Score 93;  DB 44;  Length 348;
Best Local Similarity  62.6%;  Pred. No. 5.8e-17;
Matches 161;  Conservative  0;  Mismatches 95;  Indels  1;  Gaps  1;

Qy    32  acacaccatcaaggacttacaggtgtcgcaatgtgacacagtcaaaacaattggc   91
          ||| ||| ||| ||||| || || ||| ||| || | | |||||||||| |||||||
Db    91  ACATGCCTATCTACATGTACTCGGTGTGCAACGTGGTATCCGGCGACCAGTCAACTGGC  150

Qy    92  tgagaacaaactgggt-cccaggggaactcagctcagaagaagattttatgtggagctcaagttc  150
          ||  | | ||| ||  | ||| ||||| ||| |  | | | ||| | || ||||| ||||||||
Db   151  TCCGCCAACTGGTTGTGTACCGGGAGGAGGCCTGACGCCGCATCTATTGAACTCAAGTTC  210

Qy   151  actctacgagagtcgcaatagacattccattggtgttttaggagagacattcaac  210
          ||| |  | | | || ||| ||  ||||| |||||  |||||| |||||| ||||
Db   211  ACGGAGCGGAGACTGTAACACGCTTCTCGAAGCGCCCAGCTCAATGCATAGAGACCTTCAAC  270

Qy   211  ctgtactacatggagctgtgatgatcatgggtcaaatttcgaggagacattcaca  270
          || |||||| | ||| | |||| | |  ||||  ||||| ||| |||| |||||||
Db   271  CTCTACTACTATGCAGAGACGAGTGGACTATGGCACCAACTTCCAGAAGCGCCAGTTCACC  330

Qy   271  aagattgacacattgc  287
          |||| |||||| | ||
Db   331  AAGAGTTGACACCATCGC  347
```

Qy   364 cctgtcaacaagagaggattttattgtggtcattcaagatgttggtcgttgcccttg 423
          |||| ||| || | ||||| ||||| || ||||||| || |||||| ||
Db   311 CCTGTGTCCCGCACGGCGTTCTACCTGGCCTTCTACTGTGTAACGGGGGACGGCGGAGTGGGCTG 370

Qy   424 gtgtc 428
          |||||
Db   371 GTGCC 375

RESULT 13
AI967980/c
LOCUS   AI967980   418 bp   mRNA   EST   25-AUG-1999
DEFINITION   vu11f04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2516671 3'
             similar to SW:EPA8_MOUSE O09127 EPHRIN TYPE-A RECEPTOR 8 PRECURSOR
             ; mRNA sequence.
ACCESSION   AI967980
VERSION   AI967980.1   GI:5764798
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 418)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3189146.
             Contact: Robert Strausberg, Ph.D.
             Email: Robert_Strausberg@nih.gov
             Tel: (301) 496-1550
             Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
             R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
             Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
FEATURES   Location/Qualifiers
     source   1..418
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2516671"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIds:1257096-1258631,
                     169064-1470083, and 1475592-1467743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT   152 c   116 g   77 t
ORIGIN
        73 a

Seq primer: -40UP from Gibco.
Best Local Similarity 13.4%;   Score 88.2;   DB 63;   Length 418;
Matches 154; Conservative   0; Mismatches 88;   Indels   3;   Gaps   1;

Qy   400 gatgttggttggctggtcctggtgtcctggttgagagtatacttcaaaaagtgccattt 459
          | | ||| || ||| ||| | | | || ||| ||| || ||||| | ||
Db   417 GACATAGGTGCCTCGCCTGGCCATCCTCTCTCCGCATCTTACTATPAAAGAAGTGCCCTGCC 358

Qy   460 acagtgaagaaatctggctctatgttccagacacggtaccatggactccccagtcc---ctg 516
          ||| ||| ||||||||| || || ||| | || ||| |||||| ||| |||
Db   516

Db   357 ATGGTGCGGCAAATCTGGCTGCCTTCTCGGAGGCAGCAGTGACGCGGGGGCCGACTCGTCCCTCACTG 298
          |||| || ||| ||||| || || || | ||| ||| || ||| |||| || || ||| ||
Qy   517 gtggaggttagagggtcttgtgtcaacaaatctaaggaggaaagatccctcaaggatgtac 576

Db   297 GTGGAGGTGAGGGGCCAGTGCGTGCGGGCACTCAGAGGACANACACCCAAGATGTAC 238
          ||| || |||| | |||| | ||| || || ||||| || || |||||||||
Qy   577 tgcagtaacacagaaggggaatggccttgtaccccattgacaagttcctcgacctgctggctat 636

Db   237 TGCAGCGCCGGAGGGCCGGCCAGTGGCTCGTGCCCATCAGGCGAAATCGCTGTGCGAGTGCCGGCTAC 178
          |||| |||| ||| ||||||| | || ||| || | |||| ||| ||| || |||| ||
Qy   637 gaaga 641
          |||||
Db   177 GAGGA 173

RESULT 14
AW366937/c
LOCUS   AW366937   344 bp   mRNA   EST   04-FEB-2000
DEFINITION   IL0-HT0156-251099-132-a12 HT0156 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW366937
VERSION   AW366937.1   GI:6871587
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 344)
AUTHORS   HCGP http://www.ludwig.org.br/ORSTES.
TITLE   The FAPESP/LICR Human Cancer Genome Project
JOURNAL   Unpublished (1999)
COMMENT   On Mar 16, 1998 this sequence version replaced gi:2961824.
             Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?ti-IL06t2-ILO-HT0156-
             251099-132-a12&t3-1999-10-25&t4-1)
             Seq primer: puc 18 forward
             High quality sequence start: 30
             High quality sequence stop: 343.
FEATURES   Location/Qualifiers
     source   1..344
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0156"
                     /clone="HT0156"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT   73 a   98 c   95 g   78 t
ORIGIN

Query Match   13.0%;   Score 86;   DB 80;   Length 344;
Best Local Similarity 60.1%; Pred. No. 6.6e-15; Mismatches 95;   Indels   0;   Gaps   0;
Matches 143; Conservative   0;

Qy   29 attacacaccccatcaggacttaccaggttgcaatgtcatggaccacagtcaaaacaatt 88
          || | ||| ||||||| |||| |||||||| | ||||| || |||||| || ||||
Db   268 ATGACATGCCGATCNACATGTACTACCGTGTGCAACGTGAATGTCTGGCGACGAGGACAACT 209

Qy   89 ggctgagaacaaactgggtccccaggaacctcagctcaagaagatttatgtggagctcaagt 148
          ||| || ||| ||| |||||||| |||||| ||||||||| ||| |||||||||||
Db

RESULT 15

```
Db    208  GGCGGCCGCACCAACTGGGTGTGTACCGAGGAGGACGTGAGCGTATCTTCATTGAGCTCAAGT
              ||||||   ||| | ||||||  ||||||  ||||   |||||||  || ||||||||
Qy    149  tcacctactacagaagactgccaatagccattccattggtttaggaacttgcaagagagacattca 149

Db    148  TTACTGTACGTGACTGCAACAGCTTCCCTGGTGGCGCCCAGCTCCTGCAAGGAGACTTTCA 89
              ||| |||| ||||||  ||||||| |  || |||||| ||| || ||||||| |||
Qy    209  acctgtactacatggagtctgatgatgatcatgtgggtgaaatttcgagagcatcagtt 266

Db    88   ACCTAACTACTCCGGAGTCACGGACCTGGACTACGGCTCCAATTCCAGAAGCGCCCTGTT 31
              |||||| ||| ||||||| |||  |||||||       ||| |||||
```

Query Match   11.5%;   Score 75.6; DB 64;   Length 314;
Best Local Similarity   62.6%;   Pred. No. 7.4e-12;
Matches 117;   Conservative   0;   Mismatches 70;   Indels   0;   Gaps   0;

```
Qy    32   acacacccatcaggacttaccaggttgtgcaatgtcatggaccacagtcaaaacaattggc 91
              |||  ||  |||||| ||  || | ||||||| ||||||||||||||| ||| | ||||
Db    124  ACATGCCTTATTCTACATGTTACTCGGTGTTGCCCAAGTGTGGTATTCCGGCGACCAGGACAAACTGGC 183

Qy    92   tgagaacaaactggggtcccccagggaactcagctcagaagagatttatgtggagctcaagttca 151
              || ||||| ||||  || ||||||| ||||| ||||  || |||| ||| | |||||||||||
Db    184  TCCGCCACCAACTGGGTGTGTACCGAGGAGGACGCGAGCGCCATCTTTATTGAGCTCAAGTTCA 243

Qy    152  ctctacgagactgcaatagcattccattggtctttaggaacttgcaaggagagacattcaacc 211
              ||| ||||||||| || || |||| || |||  || ||||| || |||||||| ||||| ||
Db    244  CGGTGCGAGACTGTAACAGCTTCCCTGGTGGCGCCCAGCTCCTCATGCANAGAGACCTTCAAC 303

Qy    212  tgtacta 218
              | ||||
Db    304  TCTACTA 310
```

Search completed: May 15, 2000, 11:20:21
Job time: 19144 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 12:05:06 ; Search time 102.62 Seconds
                       (without alignments)
                       816.623 Million cell updates/sec

Title: US-09-104-340-8
Perfect score: 660
Sequence: 1 tgggaagaagatcagtggtgt..........aaagaggtttatgtgccaa 660
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 226296 seqs, 63486255 residues

Total number of hits satisfying chosen parameters: 452592

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
                4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
                5: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
                6: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                7: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	3132	1	US-08-167-919A-9	Sequence 9, Appl
2	660	100.0	3132	5	US-08-715-106-9	Sequence 9, Appl
3	463.2	70.2	3254	1	US-08-162-809-15	Sequence 15, Appl
4	327	49.5	3116	4	US-08-449-645A-14	Sequence 14, Appl
5	327	49.5	3116	4	US-08-702-367A-14	Sequence 14, Appl
6	327	49.5	3116	6	PCT-US95-04681-14	Sequence 14, Appl
7	327	49.5	3348	1	US-08-222-616-34	Sequence 34, Appl
8	327	49.5	3348	6	PCT-US95-04228-34	Sequence 34, Appl
9	321.2	48.7	3592	3	US-08-469-537A-100	Sequence 100, Appl
10	317.4	48.1	3906	3	US-08-469-537A-102	Sequence 102, App
11	317.4	48.1	4165	2	US-08-442-248-1	Sequence 1, Appl
12	317.4	48.1	4165	2	US-08-440-815-1	Sequence 1, Appl
13	317.4	48.1	4322	2	US-08-673-789-1	Sequence 1, Appl
14	303	45.9	3162	4	US-08-449-645A-12	Sequence 12, Appl
15	303	45.9	3162	4	US-08-702-367A-12	Sequence 12, Appl
16	303	45.9	3162	6	PCT-US95-04681-12	Sequence 12, Appl
17	303	45.9	4529	4	US-08-449-645A-16	Sequence 16, Appl
18	303	45.9	4529	4	US-08-702-367A-16	Sequence 16, Appl
19	303	45.9	4529	6	PCT-US95-04681-16	Sequence 16, Appl
20	298.2	45.2	2323	6	PCT-US96-00419-6	Sequence 6, Appl
21	298.2	45.2	2901	5	PCT-US96-00419-4	Sequence 4, Appl
22	298.2	45.2	4304	6	PCT-US96-00419-1	Sequence 1, Appl
23	245.2	37.2	3056	1	US-08-162-809-21	Sequence 21, Appl
24	245.2	37.2	3056	1	US-08-162-809-3	Sequence 3, Appl
25	245.2	37.2	3125	1	US-08-162-809-19	Sequence 19, Appl
26	219.6	33.3	2820	1	US-08-162-809-5	Sequence 5, Appl
27	186.8	28.3	4049	1	US-08-162-809-17	Sequence 17, Appl
28	186.8	28.3	4097	1	US-08-162-809-11	Sequence 11, Appl
29	168	25.5	2962	4	US-08-449-645A-10	Sequence 10, Appl
30	168	25.5	2962	4	US-08-702-367A-10	Sequence 10, Appl
31	168	25.5	2962	6	PCT-US95-04681-10	Sequence 10, Appl
32	138.4	21.0	3546	1	US-08-162-809-9	Sequence 9, Appl
33	138.4	21.0	3591	1	US-08-162-809-13	Sequence 13, Appl
34	119.8	18.2	3776	1	US-08-162-809-7	Sequence 7, Appl
35	117.6	17.8	2982	2	US-08-348-143-2	Sequence 2, Appl
36	117.6	17.8	2982	2	US-08-348-143-3	Sequence 3, Appl
37	117.6	17.8	4027	1	US-08-571-785-4	Sequence 4, Appl
38	117.6	17.8	4027	2	US-08-348-143-4	Sequence 4, Appl
39	117.6	17.8	4027	2	US-08-571-785-3	Sequence 3, Appl
40	98	14.8	3133	1	US-08-162-809-1	Sequence 1, Appl
41	79.8	12.1	3969	1	US-08-436-044-5	Sequence 5, Appl
42	79.8	12.1	3969	1	US-08-222-616-23	Sequence 23, Appl
43	79.8	12.1	3969	3	US-08-435-054-5	Sequence 5, Appl
44	79.8	12.1	3969	6	PCT-US95-04228-23	Sequence 23, Appl

ALIGNMENTS

RESULT   1
US-08-167-919A-9
Sequence 9, Application US/08167919A
Patent No. 5674691.
GENERAL INFORMATION:
;   APPLICANT: Boyd, Andrew W.
;   APPLICANT: Simpson, Richard J.
;   APPLICANT: Wicks, Ian
;   APPLICANT: Ward, Larry D.
;   APPLICANT: Wilkinson, David
;   TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
;   TITLE OF INVENTION: AND USE THEREOF
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Scully, Scott, Murphy & Presser
;   STREET: 400 Garden City Plaza
;   CITY: Garden City
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 11530
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/167,919A
;   FILING DATE: 18-APR-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PK6841 (AU)
;   FILING DATE: 21-JUN-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PK9992 (AU)
;   FILING DATE: 12-DEC-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/AU92/00294
;   FILING DATE: 19-JUN-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Digilio, Frank S.
;   REGISTRATION NUMBER: 31,346
;   REFERENCE/DOCKET NUMBER: 9159
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (516) 742-4343
;   TELEFAX: (516) 742-4366
;   TELEX: 230 901 SANS UR
;   INFORMATION FOR SEQ ID NO: 9:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3132 base pairs
;   TYPE: nucleic acid

STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,106
FILING DATE: 18-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,919
FILING DATE: 18-APR-1994
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3132 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 100..3048
US-08-715-106-9

Query Match   100.0%;   Score 660;   DB 5;   Length 3132;
Best Local Similarity   100.0%;   Pred. No. 8.7e-208;
Matches 660;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 tgggaagagatcagtggtggtggtgatgaacattacacacccatcaaggacttaccaggtgtgc   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 TGGGAAGAGATCAGTGGTGGTGGTGATGAACATTACACACCCATCAAGGACTTACCAGGTGTGC 312

Qy   61 aatgtcatgaccacagtcaaaaacaaattggctgagaacaaacaaactgggtccccaggaactca 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 AATGTCATGGACCACAGTCAAAACAATTGGCTGAGAACAAACTGGGTCCCCAGGAACTCA 372

Qy   121 gctcagaagagattatgtggagctcaagttcactctaacggagactgcaatagcattccatg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 GCTCAGAAGAGATTTATGTGGAGCTCAAGTTCACTCTACGGAGACTGCAATAGCATTCCATTG 432

Qy   181 gtttttaggaacttgcaaggagagacattcaaacctgtactacatggagtctgatgatgatcat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 GTTTTAGGAACTTGCAAGGAGAGACATTCAAACCTGTACTACATGGAGTCTGATGATGATCAT 492

Qy   241 ggggtgaaatttcgagagcatcagtttacacaagagattgacacactggagctcaacactgagagta 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 GGGGTGAAATTTCGAGAGCATCAGTTTACACAAGAGATTGACACCATTGCAGCTCAACACTGAGAAAGT 552

Qy   301 ttcactcaaatggatctctggggaccgtattctgaagctcaacactgagaattagagagagta 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 TTCACTCAAATGGATCTTGGGGACCGTATTCTGAAGCTCAACACTGAGATTAGAGAGAGTA 612

RESULT 2
US-08-715-106-9
Sequence 9, Application US/08715106
Patent No. 6020306
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser

—

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 100..3048
US-08-167-919A-9

Query Match   100.0%;   Score 660;   DB 1;   Length 3132;
Best Local Similarity   100.0%;   Pred. No. 8.7e-208;
Matches 660;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 tgggaagagatcagtggtggtggtgatgaacattacacacccatcaaggacttaccaggtgtgc   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 TGGGAAGAGATCAGTGGTGGTGGTGATGAACATTACACACCCATCAAGGACTTACCAGGTGTGC 312

Qy   61 aatgtcatgaccacagtcaaaaacaaattggctgagaacaaacaaactgggtccccaggaactca 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 AATGTCATGGACCACAGTCAAAACAATTGGCTGAGAACAAACTGGGTCCCCAGGAACTCA 372

Qy   121 gctcagaagagattatgtggagctcaagttcactctaacggagactgcaatagcattccatg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 GCTCAGAAGAGATTTATGTGGAGCTCAAGTTCACTCTACGGAGACTGCAATAGCATTCCATTG 432

Qy   181 gtttttaggaacttgcaaggagagacattcaaacctgtactacatggagtctgatgatgatcat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 GTTTTAGGAACTTGCAAGGAGAGACATTCAAACCTGTACTACATGGAGTCTGATGATGATCAT 492

Qy   241 ggggtgaaatttcgagagcatcagtttacacaagagattgacacactggagctcaacactgagagta 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 GGGGTGAAATTTCGAGAGCATCAGTTTACACAAGAGATTGACACCATTGCAGCTCAACACTGAGAAAGT 552

Qy   301 ttcactcaaatggatctctggggaccgtattctgaagctcaacactgagaattagagagagta 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 TTCACTCAAATGGATCTTGGGGACCGTATTCTGAAGCTCAACACTGAGATTAGAGAGAGTA 612

Qy   361 ggtccctgctcacaagagaaggggattttattttgtggttgttgttgttgcc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   613 GGTCCTGTCAACAAGAGAAGGGGATTTTATTTGGCATTTTCAAGATGTTGGTGTTGTTGCC 672

Qy   421 ttggtgtctgtgagagtatatactt caaaaagtgcccattcaagtgaagaatctggctatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673 TTGGTGTGTCTGTGAGAGTATACTTCAAAAAGTGCCCATTTACAGTGAAGAATCTGGCTATG 732

Qy   481 tttccagacacggtaccccagtgactcccagtccctggttgggaagttagagggtctgtgtc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   733 TTTCCAGACACGGTACCCCATGGACTCCCAGTCCCTGTGGGAGGTTAGAGGGTCTTGTGTC 792

Qy   541 aacaattcctaggagggaagatcctccaaggatgtactgcagtacagaagaggcgaatggctt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   793 AACAATTCTAAGGAGGAAGATCCTCCAAGGATGTACTGCAGTACAGAAGGCGAAATGGCTT 852

Qy   601 gtacccaccattgtcaagtgttcctgcaatgctgtctggctatgaagaaaagaggtttatgtgtgccaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   853 GTACCCACCATTGGCAAGTGTTCCTGCTATGATGAAGAAAAGAGGTTTTAGGTGCCAA 912

```
Qy    361 ggtcctgtcaacaagaagggaggattttatttgggcattcaagatgttggttgttgttgcc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    613 GGTCCTGTCAACAAGAAGGGAGGATTTTATTTGGGCATTCAAGATGTTGGTTGTTGTTGCC 672

Qy    421 ttggtgtctgtgagagtatacttcaaaaagtgcccattacatcagtgaagaatctggctatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    673 TTGGTGTCTGTGTGAGAGTATACTTCAAAAAGTGCCCATTACAGTGAAGAATCTGGCTATG 732

Qy    481 tttccagacacggtacccatggaactccccgtcccctggtggaaggttagagggtcttgtgtc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 TTTCCAGACACGGTACCCATGGAACTCCCCAGTCCCCTGGTGGAAGGTTAGAGGGTCTTGTGTC 792

Qy    541 aacaattctaaggaggagaaatcctccaaggacgatgtactgcagtacacagaggccgaatggctt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 AACAATTCTAAGGAGGAGAAATCCTCCAAGGACGATGTACTGCAGTACACAGAGGCCGAATGGCTT 852

Qy    601 gtacccattggcaagtgttcctgcaatgctgtcatgagaagaagagcaatggctt 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    853 GTACCCATTGGCAAGTGTTCCTGCAATGCTGTCATGAGAAGAAAGAGGTTTTATGTGCCAA 912
```

RESULT   3
US-08-162-809-15
; Sequence 15, Application US/08162809
; Patent No. 5457048
;   GENERAL INFORMATION:
;   APPLICANT: Pasquale, Elena B.
;   APPLICANT: Sajjadi, Fereydoun G.
;   TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
;   TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
;   NUMBER OF SEQUENCES: 26
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: CAMPBELL AND FLORES
;   STREET: 4370 La Jolla Village Drive, Suite 700
;   CITY: San Diego
;   STATE: California
;   COUNTRY: United States of America
;   ZIP: 92122
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/162,809
;   FILING DATE:
;   CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Campbell, Cathryn A.
;   REGISTRATION NUMBER: 31,815
;   REFERENCE/DOCKET NUMBER: P-LJ 9503
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 535-9001
;   TELEFAX: (619) 535-8949
;   INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3254 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 32..2980

US-08-162-809-15

Query Match   70.2%;   Score 463.2; DB 1;   Length 3254;
Best Local Similarity   81.4%;   Pred. No. 6.3e-143;
Matches 537; Conservative   0; Mismatches 123;   Indels   0;   Gaps   0;

```
Qy      1 tgggaggagatcagtggtgtggatgaacattacacacccatcaggacttaccaggtgtgc 60
          ||||||||||||||| ||||| ||||| || |||||| || ||| ||||| |||| ||
```

```
Db    182 TGGGAAGAGATTAGTGGTGGTTGATGAGGCATTATACTCCAATCAGAACAACTTACCAAGAGAGC 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aatgtcatggaccacagtcaaaaacaattggctgagaacaaactgggtcccaggaacta 120

Db    242 AATGTTATGGATCACAGTCAAAAACAATTGGCTGAGAACAAATTCCACGCAATTCA 301

Qy    121 gctcagaaagattcatgtgggagctcaagttcactcacgaactgtcaaatagcaattcattg 180
          |||||||| ||||| ||||||| ||||||| || |||||||||| || |||||| ||||
Db    302 GCGCAGAAGAATATATGTGGGAGCTCAAGTTCACCTTGAGGGACTGCAATAGTATCCCTCTA 361

Qy    181 gtttttaggaacttgcaaggagacattcaacctgtactacatggagtctgatgatgatcat 240
          ||| || ||||||| ||||| || ||||| || || |||||| ||| |||| || ||
Db    362 GTTCTCGGGCACTTGCAAAGTTGCAAATTTCAAATCGTGATTACACAGGAGTCTGAATGAGAGC 421

Qy    241 ggggtgaaaattcgagagagcatcagtttacaaagagttacacaccattgcagtctgataagt 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 TTGGCAAAGTTCAGAGAGCACCAAATTTACGAAGAGATTGACACCATGGCCGGCTGAATGAGAGC 481

Qy    301 ttcacctcaaatggatcttttggggaacgtattctgaagctcaacactgagattagaaagta 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 TTCACCCAGATCGGATCTTTGGGGGACCCGGATTTCTCAACGTGAATACCGAAGTCCGCGAGGTG 541

Qy    361 ggtcctgtcaacaagaaggggattttatttgggcattcaagatgttggttgttgttgcc 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 GGACCGGTTAGTAAGAAGGGCTTTTACTTGGCTTCCAAGATGTAGAGGTGCGATGTTGTTGGC 601

Qy    421 ttggtgtctgtgagagtatacttcaaaaagtgcccattacagtgaagaatctggctatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 TTAGTCTCGGGTGCGAGTGTACTTCAGAGAAGTGCCCTTTCACTGTCAAGAACCTCGGCCATG 661

Qy    481 tttccagacacggtacccatggaactcccggtgtgtggaaggttagagggtcttgtgtc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 TTTCCAGATACAGTTCCTATGGAGTCCCAGTCCCTGTGGAAGGTGCGGGGTTCTTGTGTC 721

Qy    541 aacaattctaaggaggagaatccctccaaggatgtactgcagtacacagaggccgaatggctt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 AATCAATTCCAAGGAGGAGAGAGCCACCAAGAATGTAGTGACTGCAGTGAGTGTTTGGCTA 781

Qy    601 gtacccattggcaagtgttcctgcaatgctggctatgaagggaagagagggtttatgtgccaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782 GTGCCCATAGGGAAGTGCTTGTGTAAAGCTGGCTATGAAGAGAAGGAGGCTTTGCGTGCCAA 841
```

RESULT   4
US-08-449-645A-14
; Sequence 14, Application US/08449645A
; Patent No. 5981245
;   GENERAL INFORMATION:
;   APPLICANT: Fox, Gary M.
;   TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
;   TITLE OF INVENTION: Kinases
;   NUMBER OF SEQUENCES: 43
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Amgen Patent Operations/RBW
;   STREET: 1840 Dehavilland Drive
;   CITY: Thousand Oaks
;   STATE: California
;   COUNTRY: USA
;   ZIP: 91320
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/449,645
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Winter, Robert B.
;   REFERENCE/DOCKET NUMBER: A-287
;   INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:

Left column

```
;  LENGTH: 3116 base pairs
;  TYPE: nucleic acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: cDNA
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: 34..2994
US-08-449-645A-14

Query Match             49.5%;  Score 327;  DB 4;  Length 3116;
Best Local Similarity   69.4%;
Matches  460;  Conservative    0;  Mismatches 200;  Indels  3;  Gaps  1;
                               Pred. No. 4.7e-98;

Qy     1 tgggaagagatcagtcggtgtggatgaacattacacaccatcaggaactaccaggtgtgc   60
         ||||| ||  ||  |  |  | |  | ||  || ||| ||| |   || |||| ||||||
Db   193 TGGGAGGAAGTGAGTATCATGGATGAAAAAATACCACCAATCCGAACCTACCAAGTGTGC  252

Qy    61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactgggtccccaggaactca  120
         |||||  | ||||  | |   |   | ||||| || ||  |  |||| | |||| |||||
Db   253 AATGTGAATGGAACCAGCCAGAATAACTGGCTACGAACTGATTGGGATCACCCGAGAGGG  312

Qy   121 gctcagaagatttatgtgggagctcaagttcactcacgagactgcaatagcattccattg  180
         |||||| | |||||| |  || |||   |||||   | |||||||||||| ||||| |||
Db   313 GCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC  372

Qy   181 gtttaggaacttgcaaggaggaacattcaacctgtactacatggagtctgatgatcat    240
         || |||||| |||||||||  |  | ||  ||| |||| | |||| | |||| ||| |
Db   373 GTCATGGGGACTTGCAAGGAGACGTTTAACCTGCTACTACTATGAATCAGACAACGACAAA  432

Qy   241 ggggtgaaattcgagagcatcagttacaaagaattgacaccatcgcagctgatgaaagt   300
         |||||| | |||||| | |||| |  ||  || |  |||||  | ||||||||||| ||
Db   433 GAGCGTTTCATCAGAGAGAACCAGTTTGTCAAAATTGACACCATTGCTGCTGATGAGAGC  492

Qy   301 ttcactcaaatggatcttctggggaccgtattcctgaagctcaacacttagaagaagta   360
         |||| || |||||| |  |||||| | |  |  |  || |||||||  || |||||| |
Db   493 TTCACCCAAGTGACATTGGTCGACATTGGTGTTCTAAAAAGTGTCCACTCACAGTCCCAAC  552

Qy   361 ggtcctgtcaacaagaaggagggatttattttggcattcaagatgtttggtgcttgtgcc  420
         || || || || || ||| || | ||| |  ||||| ||||| |  | || | |  |||
Db   553 GGGGCATTAACCAAAAAGGGGTTTTACCTGCCTTTCAAGGATGGGGGCCTTGACCTGCATG  612

Qy   421 ttggtgtctgtgagagtatactccaaagaattgacaccatcgcagctgatgaagaagta   480
         || || || ||| | |  |  | ||  || |  |||||  | ||||||||||  ||||
Db   613 CTGGTATCAGTCCGTGTTCTATAAAAAGTGTCCACTCACAGTCCCAAATCTGGGCCCAG   672

Qy   481 tttccagaacacggtaccccatg---gactccagtccctggtggaggttagagggtcttgt  537
         || ||| | |||| | | ||| |   |||| || | | ||||||| || |||| | ||||
Db   673 TTTCCTGACACCATCACAGGGGCTGATAACGGTCTTCCCTGGTGGAAGTTCGAGGCTCCTGT  732

Qy   538 gtcaacaattctaaggaggaagatcctccaaggatgtactcagtacagaaggcgaatgg   597
         || || |||| |||||| || || || ||||| |||| ||| ||||| || | ||| ||
Db   733 GTCAACAACTCAGAGAGAAGAAATGTCCAAAAATGTACTGTGGGGCCAGATGAGGTGAATGG  792

Qy   598 cttgtgtaccccattggcaagtgctggctatgatgaagaaagaggttttatgtgtc     657
         || ||| | ||||||||| || ||| | | |||||| | | ||| || |||| |
Db   793 CTGGTACCAATGGCAACTGCCTATGGGGCATGAGGAGCGGGGAGAATGC              852

Qy   658 caa  660
         |||
Db   853 CAA  855
```

Result block

```
RESULT 5
US-08-702-367A-14
Sequence 14, Application US/08702367A
Patent No. 5981246

GENERAL INFORMATION:
APPLICANT:  Fox, Gary M.
TITLE OF INVENTION:  EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION:  Kinases
NUMBER OF SEQUENCES:  43
```

Right column

```
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Amgen Patent Operations/RBW
;     STREET: 1840 Dehavilland Drive
;     CITY: Thousand Oaks
;     STATE: California
;     COUNTRY: USA
;     ZIP: 91320
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/702,367A
;     FILING DATE:
;     CLASSIFICATION:
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Winter, Robert B.
;     REFERENCE/DOCKET NUMBER: A-287
;  INFORMATION FOR SEQ ID NO: 14:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 3116 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;        NAME/KEY: CDS
;        LOCATION: 34..2994
US-08-702-367A-14

Query Match             49.5%;  Score 327;  DB 4;  Length 3116;
Best Local Similarity   69.4%;
Matches  460;  Conservative    0;  Mismatches 200;  Indels  3;  Gaps  1;
                               Pred. No. 4.7e-98;

Qy     1 tgggaagagatcagtcggtgtggatgaacattacacaccatcaggaactaccaggtgtgc   60
         ||||| ||  ||  |  |  | |  | ||  || ||| ||| |   || |||| ||||||
Db   193 TGGGAGGAAGTGAGTATCATGGATGAAAAAATACCACCAATCCGAACCTACCAAGTGTGC  252

Qy    61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactgggtccccaggaactca  120
Db   253 AATGTGAATGGAACCAGCCAGAATAACTGGCTACGAACTGATTGGGATCACCCGAGAGGG  312

Qy   121 gctcagaagatttatgtgggagctcaagttcactcacgagactgcaatagcattccattg  180
Db   313 GCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC  372

Qy   181 gtttaggaacttgcaaggaggaacattcaacctgtactacatggagtctgatgatcat    240
Db   373 GTCATGGGGACTTGCAAGGAGACGTTTAACCTGCTACTACTATGAATCAGACAACGACAAA  432

Qy   241 ggggtgaaattcgagagcatcagttacaaagaattgacaccatcgcagctgatgaaagt   300
Db   433 GAGCGTTTCATCAGAGAGAACCAGTTTGTCAAAATTGACACCATTGCTGCTGATGAGAGC  492

Qy   301 ttcactcaaatggatcttctggggaccgtattcctgaagctcaacacttagaagaagta   360
Db   493 TTCACCCAAGTGACATTGGTCGACATTGGTGTTCTAAAAAGTGTCCACTCACAGTCCCAAC  552

Qy   361 ggtcctgtcaacaagaaggagggatttattttggcattcaagatgtttggtgcttgtgcc  420
Db   553 GGGGCATTAACCAAAAAGGGGTTTTACCTGCCTTTCAAGGATGGGGGCCTTGACCTGCATG  612

Qy   421 tggtgtctgtgagagtatactccaaagaattgacaccatcgcagctgatgaagaagtatg  480
Db   613 CTGGTATCAGTCCGTGTTCTATAAAAAGTGTCCACTCACAGTCCCAAATCTGGGCCCAG   672

Qy   481 tttccagacacggtaccccatg---gactccagtccctggtggaggttagagggtcttgt  537
Db   673 TTTCCTGACACCATCACAGGGGCTGATAACGGTCTTCCCTGGTGGAAGTTCGAGGCTCCTGT  732

Qy   538 gtcaacaattctaaggaggaagatcctccaaggatgtactcagtacagaaggcgaatgg   597
```

RESULT 6
PCT-US95-04681-14
; Sequence 14, Application PCT/US9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..2994
PCT-US95-04681-14

Query Match   49.58;   Score 327; DB 6; Length 3116;
Best Local Similarity 69.4%;   Pred. No. 4.7e-98;
Matches 460; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

Qy   733 GTCAACAACTCAGAAGAGAAGAGAAAATGTACCTGTGGGGCAGATGGTGAATGG 792
Qy   598 cttgtacccattgggcaagtcgtcctgcaatgctggctatgaagaagaggtttatgtgtc 657
Db   793 CTGGTACCCATTGGCAACTGCCTATGCAACGCTGGGCATGAGGAGGCGGGGAAATGC 852
Qy   658 caa 660
Db   853 CAA 855

Qy   1 tgggaagatcagtggtgtggatgaacaattacacacccatcaggacttacaggtggtgc   60
Db   193 TGGGAGGAGTCAGTATCATGGATGAAAAAATACACCAATCCGAACCTACCAAGTGTGC   252
Qy   61 aatgtcatggaccacagtcaaaacaattgggtcggagaacaaactggtccccaggaactca 120
Db   253 AAATGTGAATGGAACCACCAGCAGAATAACTGGCTACGAACTGATTGGATCACCCGAGAAGGG 312
Qy   121 gctccagaagatttatgtggagctcaagttcactcctcacgaagcatagcaattccattg 180
Db   313 GCTCCAGGGGTGTATATTGAGATTAAAATTCACCTTGAGGGACTGGACTGCAATAGTCTTCCGGC 372
Qy   181 gttttaggaacttgcaaggaggacaattcaacctgtactacatggagtcctgatgatcat 240
Db   373 GTCATGGGGACTTGCAAGGAGGACGTTTAACCTGTACTACTATGAATCAGACAACGACAAA 432
Qy   241 ggggtgaaattcgagagcatcagttacacaaagagattgacacaccattgcagctgatgaaagt 300

RESULT 7
US-08-222-616-34
; Sequence 34, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935
; FILING DATE: 22-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 821P2
; TELECOMMUNICATION INFORMATION:

Db   433 GAGGGTTCATCAGAGAGAACCAGTTTGTCAAAATTGACACCCATTGCTGCTGATGAGAGC 492
Qy   301 ttcactcaaatgtgatcttgggagaccgtattctgaagtcaacactgagaattagaagta 360
Db   493 TTCACCCAAGTGGACATTGGTGACAGAATCATGAAGCTGCAACACCGAGATGGTA 552
Qy   361 ggtcctgtcaacaagaagggatttattttggcatttcaagagatgttgggtgcttgtgttgcc 420
Db   553 GGGCCATTAAGCAAAAAAGGGTTTTTACCTGGCTTTTCAGGATGTGGGGCCTGCCATCGCC 612
Qy   421 ttggtgtctgtgggagtatacttccaaaaaagtgccatttacagtgaagaatctggctatg 480
Db   613 CTGGTATCAGTCCGTGTGTTCTATAAAAAGTGTCCACTCACAGTCCGGCCAAATCTGGGCCCAG 672
Qy   481 tttcccagacacggtacccatg--gactcccagtccctggttgagggttagagggtcttgt 537
Db   673 TTTCCTGACACCATCACACAGGGGCGTGATACGTCTTCCCGTGTGAAGTTCGAGCGTCCGT 732
Qy   538 gtcaacaattctaaggaggaagaagatcctccaaggagtgtactgcagtacagagaaggcgaatgg 597
Db   733 GTCAACAACTCAGAAGAGAAGAGAAAATGTACCTGTGGGGCAGATGGTGAATGG 792
Qy   598 cttgtacccattgggcaagtcgtcctgcaatgctggctatgaagaagaggtttatgtgtc 657
Db   793 CTGGTACCCATTGGCAACTGCCTATGCAACGCTGGGCATGAGGAGGCGGGGAAATGC 852
Qy   658 caa 660
Db   853 CAA 855

; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3348 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-222-616-34

Query Match   49.5%;   Score 327;   DB 1;   Length 3348;
Best Local Similarity   69.4%;   Pred. No. 4.9e-98;
Matches   460;   Conservative   0;   Mismatches 200;   Indels   3;   Gaps   1;

```
Qy    1 tgggaagagagatcagtggtgtggatgaacattaacacaccatcaggacttaccaggtgtgc 60
        |||||  || || ||||| | |||||||||  |||| ||||||||||| |||| | ||||||||
Db  160 TGGGAGGAAGTGAGTGAGTCATGGATGAAAAAAAATACACCAATCCGAACCTACCAAGTGTGC 219

Qy   61 aatgtcatggaccacagtcaaaacaatggctgagaacaaactgggtccccaggaactca 120
        ||||| ||||| |  |||| ||| ||||||| |||  |||||||||| ||| ||||||||
Db  220 AATGTGATGGAACCCAGCCAGAATAACTGGCTACGAACTGATTGGATCACCCCGAGAAGG 279

Qy  121 gctcagaagattatgtggagctcaaagttcaactcactcctacgaactgcaatagcattccatg 180
        ||||||||| | |||  ||||||  ||||||||  |  |||||||||| |||| || |||||||
Db  280 GCTCAGAGGGTGTATATTGGAGTAAATTCACCTGAGGGACTGCAATAGTCTTCCGGGC 339

Qy  181 gtttaggaacttgcaaggagagacattcaacctgtactacatggagtctgatgatgatcat 240
        | || |||||| | |||||||| |||| ||||| |||||||||| ||||| ||| ||||||
Db  340 GTCATGGGGACTTGCAAGGAGACGGTTTAACCTGTACTACTACTACATGAATCAGACAACGACAAA 399

Qy  241 ggggtgaaattcgagagcatagtgttacacaagagattgacacaccattgcgcctgatgaaagt 300
        |||| |||| | |||||| | |||  |||||| |||| |  |||||||| |  |||||||  ||
Db  400 GAGCGTTTCATCAGAGAGAGAACCAGTTTGTCACAACCATTGCTGCTGATGAGAAGC 459

Qy  301 ttcaactcaaatggatcttgggaccgtattctctgaagctcaacctgagcattagagaagta 360
        |||||||  ||||| ||||| || ||  | ||||||| |||||| |  ||| | ||||||| |
Db  460 TTCACCCAAGTGGACATTGGTGACAGAATCATGAGAAGTCCGGAGTA 519

Qy  361 ggtcctgtcaacaagaaggggattttattctttggcatttcaagatgttggcttgtgttgcc 420
        |  |||||||||| ||| || |  | |||||  ||||||||||||||||| |||| ||||||
Db  520 GGGCCCATTAAGCAAAAAGGGGTTTTTACCTGGCTTTTCAGGATGTGGGGGCCTCGCCATCGCC 579

Qy  421 tggtgtctgtgagagtatactccaattacagtgaagaactggctatg 480
        |||||  |||| ||| ||||  ||||| ||||||| |  ||||||||
Db  580 CTGGTATCAGTCCGTCGTGTTCTATAAAGAAAGTGTCCACTCACAGTCCGCAATCTGGCCCAG 639

Qy  481 ttttcccagacacacggtaccccatg--gactccagtccctggttaggggttcttgt 537
        ||||| || || |  ||  ||  ||    | | |||||  | ||| ||  | |||
Db  640 TTTTCCTCACACCATCACAGGGGGGTGATACGTCTTCCCTGGTGGAAGTTCGAGGCTTCGTGT 699

Qy  538 gtcaacatttctaaggaggaagatccccaaggatgtactgcgtacaagaaggcgaatgg 597
        ||||||  |  || ||| |||| |  ||||||| ||  || |  ||  || || |||||
Db  700 GTCAACAAACTCACAGAGGAAGAGAAATGTACCACTGGTGGGCAGAATGGGTTGAAGG 759

Qy  598 cttgtaccccatggtgttcctgcaatgtctgggctcatgatgaagaaagagggtttatgtgc 657
        ||||||||| ||| ||  | || || |  |||| ||| || || ||||  |||||| ||||
Db  760 CTGGTACCCATTGGCAACTCAGAAGAGAAACAAAGTACGTGGGCAGGAGCGGAGCGGAGAATGC 819

Qy  658 caa 660
        |||
Db  820 CAA 822
```

RESULT 8
PCT-US95-04228-34
; Sequence 34, Application PC/TUS9504228
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.

Query Match   49.5%;   Score 327;   DB 6;   Length 3348;
Best Local Similarity   69.4%;   Pred. No. 4.9e-98;
Matches   460;   Conservative   0;   Mismatches 200;   Indels   3;   Gaps   1;

```
Qy    1 tgggaagagagatcagtggtgtggatgaacattaacacaccatcaggacttaccaggtgtgc 60
        |||||  || || ||||| | |||||||||  |||| ||||||||||| |||| | ||||||||
Db  160 TGGGAGGAAGTGAGTGAGTCATGGATGAAAAAAAATACACCAATCCGAACCTACCAAGTGTGC 219

Qy   61 aatgtcatggaccacagtcaaaacaatggctgagaacaaactgggtccccaggaactca 120
        ||||| ||||| |  |||| ||| ||||||| |||  |||||||||| ||| ||||||||
Db  220 AATGTGATGGAACCCAGCCAGAATAACTGGCTACGAACTGATTGGATCACCCCGAGAAGG 279

Qy  121 gctcagaagattatgtggagctcaaagttcaactcactcctacgaactgcaatagcattccatg 180
        ||||||||| | |||  ||||||  ||||||||  |  |||||||||| |||| || |||||||
Db  280 GCTCAGAGGGTGTATATTGGAGTAAATTCACCTGAGGGACTGCAATAGTCTTCCGGGC 339

Qy  181 gtttaggaacttgcaaggagagacattcaacctgtactacatggagtctgatgatgatcat 240
        | || |||||| | |||||||| |||| ||||| |||||||||| ||||| ||| ||||||
Db  340 GTCATGGGGACTTGCAAGGAGACGGTTTAACCTGTACTACTACTACATGAATCAGACAACGACAAA 399

Qy  241 ggggtgaaattcgagagcatagtgttacacaagagattgacacaccattgcgcctgatgaaagt 300
        |||| |||| | |||||| | |||  |||||| |||| |  |||||||| |  |||||||  ||
Db  400 GAGCGTTTCATCAGAGAGAGAACCAGTTTGTCACAACCATTGCTGCTGATGAGAAGC 459

Qy  301 ttcaactcaaatggatcttgggaccgtattctctgaagctcaacctgagcattagagaagta 360
        |||||||  ||||| ||||| || ||  | ||||||| |||||| |  ||| | ||||||| |
Db  460 TTCACCCAAGTGGACATTGGTGACAGAATCATGAGAAGTCCGGAGTA 519

Qy  361 ggtcctgtcaacaagaaggggattttattctttggcatttcaagatgttggcttgtgttgcc 420
        |  |||||||||| ||| || |  | |||||  ||||||||||||||||| |||| ||||||
Db  520 GGGCCCATTAAGCAAAAAGGGGTTTTTACCTGGCTTTTCAGGATGTGGGGGCCTCGCCATCGCC 579
```

```
Qy  421  ttggtgtctgtgagagtatacttcaaaaaagtgccatttacagtgaagaatctggctatg  480
         |||||||||||||||||||||||||||||||| ||||||  |  |||||||
Db  580  CTGGTTATCAGTCAGTTCCGTCGTGTTCTATAAAAAAGTGTTCCACTCACAGTCCGCAATCTGGCCCAG  639

Qy  481  tttccagacacggtacccatg---gactcccagtccctggttgagggttagagggtcttgt  537
         |||||||                 ||||| |||||||||||||||||||||||||
Db  640  TTTCCTGACACCATCACAGGGCGTCGTTTCCCTGGTTGGGAAGTTCAGGCTCCTGT  699

Qy  538  gtcaacaattctaaggagggagaagatcctccaaggatgtatctg  597
         ||||||||||||||||||||||||||||||||||||
Db  700  GTCAACAACTCAGAAGAAGAAGAGAAGAGGATGTGCCAAAAAATGTACTCGTGGGGCAGATTGGTGAAATGG  759

Qy  598  cttgtacccattggcaagtgttcctgcaatgctgctatgaagaagaggtttatgtgc  657
         |||||||||||||||||||||||||||||||||
Db  760  CTGGTACCCAFTGGCAACTGCCTATCCAACCTGGGCATGAGGAGGAGGGAGGAGAGGAGAATGC  819

Qy  658  caa  660
         |||
Db  820  CAA  822
```

RESULT 9
US-08-469-537A-100
Sequence 100, Application US/0846953 7A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisonpierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 3592 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 598..3444
FEATURE:
NAME/KEY: modified_base

LOCATION: 56
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /label= N
OTHER INFORMATION: /note= "where N = G, A, C or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3538
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /label= N
OTHER INFORMATION: /note= "where N = G, A, C or T"
US-08-469-537A-100

```
Query Match    48.7%;  Score 321.2;  DB 3;  Length 3592;
Best Local Similarity   68.9%;
Matches  456;  Conservative   0;  Mismatches  203;  Indels   3;  Gaps   1;
                                                       Pred. No. 4.1e-96;

Qy     1  tgggaagaagatcaagtgtggtgtggatggaacaataccacaccaccaagggactaccaggtagtgtgccaggaactca   60
          |||| | || ||  |||  || ||  ||| ||| || ||  ||| ||||| || ||  ||||   ||||
Db   763  TGGGATGGCCATTACTGAAATGGATGAACACAACAGCCCCATTACATACAACAGGTTAYGC  822

Qy    61  aatgtcaatggaccacagtcaaaacaattggtgctgagaacaaactgggtccccaggaactca  120
          |||||  |  |||  || | ||| || |||||  |||| || ||||  ||| ||||
Db   823  AAYTGCATTGGAACCAGAACAACTGGCTTCGTACTAACTGGACTCTCGTGAYGCT  882

Qy   121  gctcagaagaatttatgtggagctcaagttcactcactacggagactgcaataggacattcattg  180
          |||||||| |||| || ||||| ||||||||| |||| ||| ||||  || ||  ||||| ||||
Db   883  GCTCAGAAAAATCTAYGTGGAAATGAAGTTCACATTGAGAGAGATTGTAACAGCATCCCATGG  942

Qy   181  gttttaggaacttgcaaggagagacattcaaacctgtactacatggagtctgatgatgatcat  240
          || ||  |||| || ||| || ||| || || ||  ||| || |||   |||  || ||
Db   943  GTCTTGGGAACGTGTAAAGAAACATTTAACTCTGTATTATATATGGAAATCTGAGGAAATCCAC  1002

Qy   241  gggtgaaatttcgagagagcatcagtttacaaaagattgaccacaccattgcagctgatgaaagt  300
          ||  |||| | ||  ||  ||| | |  ||||| ||| || ||  ||||| |  ||||||| ||
Db  1003  GGAACTAAATTCAAGCCAAGCCAATATATAAGAATTGAACACAACAATTGCTGCGGATGAGAGT  1062

Qy   301  ttcacaccaaatgggtcatctgggaaccgtatttctgaagctcaacacacagattagagaagta  360
          || ||  || ||||| ||  || |||  |||| || ||||  || |||| |||| || |||
Db  1063  TTTACTCAGATGGATTTGGGGTGACCAAAAAATGCCCCTTCACTGAGGAACTGGGCATG  1122

Qy   361  ggtccgtgtcaacaaaggaagggatttatttggcattcaaagatgttggttgtcttgttgcc  420
          |||  | | ||||  ||| || ||| |  |||||| || ||| ||| || ||  ||||| |||
Db  1123  GGGCCAATAGAACAAGGAAGAAAGGAATTCTAYTTTGGCTTTCAAGAYATGAGCATGGCT  1182

Qy   421  ttggtgtctgtgagagtatacttcaaaaagtgcccattttacagtgaagaatctggctatg  480
          ||||| ||| |  || ||||  | |||||| ||||| || || || || |||||||||||
Db  1183  CTGGTCTCAGTCCGTGYGTTCTACAAAAAAYGCCCCTTCACTGTGCGGAAACTGGGCTATG  1242

Qy   481  tttccagacacggtaccca---tgggactcccagtcccctgggtggagggttagagggtcttgt  537
          ||||| |||||  ||  ||    || |||  | ||||| ||||| ||||| || ||| |||||
Db  1243  TTTCCTGACACCATCACAGGGCGTTGACTCTCTTCCCTGGTTGGGGAAGTTGGGAACTGGGTG  1302

Qy   538  gtcaacaattctaaggagggagaagatcctccaaggatgtatctgcagtgacagaaggcgaatgg  597
          ||   |  | || || || || ||| |  |   || | | | || || ||||| ||| |||
Db  1303  GTGAAGAGTTCTGAGGACGAGATTACTCCTAAACTACTGGACTGAGTGGAGACTGG  1362

Qy   598  cttgtacccattggcaagtgttcctgcaatgctgctatgaagaagaggtttatgtgc  657
          || || || || |||| ||| || ||||  | | | || ||| ||| |  ||||||
Db  1363  CTCGTGCCTCTTGGAAGGTGTATCTGCACTACAGGGTATGAAGAGAAATCGAGGGTTCTGC  1422

Qy   658  ca  659
          ||
Db  1423  CA  1424
```

RESULT 10
US-08-469-537A-102
Sequence 102, Application US/0846953 7A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisonpierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE

TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Regeneron Pharmaceuticals, Inc.
  STREET: 777 Old Saw Mill River Road
  CITY: Tarrytown
  STATE: NY
  COUNTRY: U.S.A.
  ZIP: 10591
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/469,537A
  FILING DATE: 06-JUN-1995
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: USSN 08/406,247
  FILING DATE: 17-MAR-1995
  APPLICATION NUMBER: USSN 08/144,992
  FILING DATE: 28-OCT-1993
  APPLICATION NUMBER: USSN 07/736,559
  FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Kempler, Ph.D., Gail M
  REGISTRATION NUMBER: 32,143
  REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 914-345-7400
  TELEFAX: 914-345-7721
  TELEX:
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
  LENGTH: 3906 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: unknown
  TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
  NAME/KEY: CDS
  LOCATION: 476..3493
US-08-469-537A-102

Query Match   48.1%;   Score 317.4;   DB 3;   Length 3906;
Best Local Similarity   68.5%;
Matches   454;   Conservative   0;   Mismatches   206;   Indels   3;   Gaps   1;

```
Qy      1 tgggaagagatcaagtggtgtggatgaacattacacacaccatcaggacttaccagtgtgc 60
            |||||  ||| |||| || |||| |||||||| |||||| ||||||||||||||||||||
Db    728 TGGGAAGAGATTGGTGAAGTTCGTGAAAACTATGCCCCCATCCACACACCTATCAAGTGTGC 787

Qy     61 aatgtcatggaccacagtcaaaacaatttggctgagaacaaactgggtccccaggaactca 120
            ||| ||||||||| |||  ||||| |||||||||||||||||| ||||| |||||||||||
Db    788 AAAGTTATGGAACAGAATCAGAATAATTGGGCTGTTGACCAGTTGGATCTCTAACGGAAGGT 847

Qy    121 gctcagaagattatgtgggagctcaagttcacctcactacgaatagcattccattg 180
            |||  | || |||||||||||||| ||||||| | || ||||| ||| |||||||
Db    848 GCTTCCAGAATTTTATTGAACTCAAGTTTACTCTGAGGGATTGCAACAGCCTTCCTGGA 907

Qy    181 gtttaggaacttgcaaggagaacattcaaactgtactacatggagtctgatgatcat 240
            ||| | || ||||| |||||| ||||| |||| |||| ||||||||| |||| || |
Db    908 GGACTGGGGACTTGCAAGGAGAACCTTTAACATGTATTATTTGAGTCGGATGATGAGAAT 967

Qy    241 ggggtgaaattcgagagcaatcaagtttacacaaagattgacacattgcagctgatgaaagt 300
            |||||||| ||||||||| |||||| ||| ||||| |||| ||||| | || ||||| |
Db    968 GGGAGAAATATCAAAGAGAACCAGTACATCACATCAAGATCATCAATTGCTGCTTGATGATGAGC 1027

Qy    301 ttcactcaaatggatcttgggacggtatttctgaagctcaacactgagagattagagaagta 360
            ||||| | ||||| ||||||| ||| |||||||||||||||||||||||||| |||||
Db   1028 TTCACCGAACTGACCTTGGAGAACCGGGTCATGAAGCTGAATACGGAGGTCAGAGAGTGTA 1087
```

Qy   361 ggtcctgtcaacaagaagggatttattggcattcaagaggattgtggctggttgtctgtgtctgtgtctgtgttgcc 420
Db 1088 GGACCTCTGAGCAAAAAGGGATTTTATCTTGCTTTCAAGAGATGTCGGGTTGCTTGCATGCCT 1147

Qy   421 ttggtgtctgtgagtgtatactccaaaaagtgcccatttacagtgaagaactcggctatg 480
Db 1208 TTCCCTGACACAGATCACTCGGACGAGATCTTCCCAGTTGCTAGAGGTGTCAGGTCCTGA 1267

Qy   481 tttccagacacggtaccccatggacttcccagtccc---tggtgaggttagaagggtctcct 537
Db 1268 GTCAAACCATTCTTGACAGACAGCGATCCTCCCAAAATTGCATTGCAGTTGCAGTGCTTGA 1327

Qy   538 gtcaacaattctaaggagaagatccctccaaaggatgtactgcagtacaagaaggcgaatgg 597
Db 1328 CTGGTTCCCATCGGGAAATTGCARTGTGCAAGGCGGGATATGAGAGGAAAAAATGGTTACCTGT 1387

Qy   658 caa 660
Db 1388 CAA 1390

RESULT 11
US-08-442-248-1
Sequence 1, Application US/08442248
Patent No. 5759863
GENERAL INFORMATION:
  APPLICANT: Caras, Ingrid W.
  APPLICANT: Winslow, John W.
  TITLE OF INVENTION: AL-1 Neurotrophic Factor
  NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
  STREET: 460 Point San Bruno Blvd
  CITY: South San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94080
COMPUTER READABLE FORM:
  MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/442,248
  FILING DATE: 15-MAY-1995
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/330128
  FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
  NAME: Torchia, Timothy E.
  REGISTRATION NUMBER: 36,700
  REFERENCE/DOCKET NUMBER: 920C4
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-8674
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
  LENGTH: 4165 bases
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
US-08-442-248-1

Query Match   48.1%;   Score 317.4;   DB 2;   Length 4165;
Best Local Similarity   68.5%;   Pred. No. 7.9e-95;

Matches   454;   Conservative   0;   Mismatches   206;   Indels   3;   Gaps   1;

```
Qy     1 tgggaagaagatcagtggtgtggatgaacattacacacccatcaggaacttaccaggtgtgc 60
          ||||||||||| |||||||||||  ||| |||| ||||||| |||||||||||||||||||
Db   793 TGGGAAGAGAGGATTGGTGVGAAGTTGVTGVGAGCTGAAGVGAAGCTATCAAGTGTGC 852

Qy    61 aatgtcatggaccacagtcaaaaacaaattgggctgagaacaaactgggtccccaggaactca 120
          |||||||||||| ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   853 AAAGTTATGGAACAGAATCAGAATAATTGGCTGTTGACCAGTGTGGATCTCTAACGAAGGT 912

Qy   121 gctcagaagattatgtggagctcaagttcacctcacgagactgcaatagcaattccattg 180
          |||  |||||| || ||||||||||||||||||   |||||||||||||| ||||||||||
Db   913 GCTTCCAGAATTTTTATTGAACTCAAGTTTTACTCTGAGGGATTGCAAACACCCTTCCTGGA 972

Qy   181 gtttaggaacttgcaaggaggagacattcaacctgtactacatggatctgatgatcatcat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   973 GGACTGGGGACTTGCAAGGAGGAGACCTTTAACATGTATAATGTATTATTTTGACTGCAGATGAGAAT 1032

Qy   241 ggggtgaaattctgagagcatcagttacaaagattgacaccattgacagtcaagctgatgaaagt 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1033 GGGAGAAATATCAAAGAGAGAACCAGTACATCAAGATACCATTGCTGCTGGATGAGAGC 1092

Qy   301 ttcactcaaatggatcttgggggaccgtattctgaagctcaacactgagattagagaagta 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1093 TTCACCGGAACTTGACCTTGGAGCGGGGTCATGAAGCTGAATACCGGAGGTCAGAGATGTA 1152

Qy   361 ggtccgtcaacagaagaaggggatttatttggcatttcaagatgttggtgcttgtgttgcc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1153 GGACCTCTGAGCAAAAAGGGGATTTTATTCTTGCTTTCCAAGATGTCGGTCGCTTGCATCGCT 1212

Qy   421 ttggtgtctgtgagagtatacttcaaaaaagtgcccattacagtgaagaatctggctatg 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1213 CTGGTTTCTGTCTGCCGTGTCTACTATAAAAATGTCCTTCGTAGTTAGACATTTGGCTGTT 1272

Qy   481 tttccagacacggtacccatggactcccagtccc---tggtgaggttagagggttcttgt 537
          ||||||||||||||||||||||||||||||||      ||||||||||||||||||||||
Db  1273 TTCCCTGACACGATCACTCGAGCAGATTCCTCCCAGTTGCTGTCTGGTGTGTCAGGCTCCTGC 1332

Qy   538 gtcaacaattctaaggaggagaatcctccaaggatgtactgcagtacagaaggcgaatgg 597
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1333 GTCAACCATTCTGTGACACAGACGATCCTCCCAAAATGCATTGCAGTACTGTGAAGGGGAGTGG 1392

Qy   598 cttgtacccattggcaagtgttcctgcagctgctataagaagaagaggtttatgtgtgc 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1393 CTGGTTCCCATCGGGAAATGCATGTGCAAGGCCGGATATGAAGAGAGAAAAATGGTACCTGT 1452

Qy   658 caa 660
          |||
Db  1453 CAA 1455
```

RESULT 12
US-08-440-815-1
Sequence 1, Application US/08440815
Patent No. 5798448
GENERAL INFORMATION:
  APPLICANT: Caras, Ingrid W.
  APPLICANT: Winslow, John W.
  TITLE OF INVENTION: AL-1 Neurotrophic Factor
  NUMBER OF SEQUENCES: 18
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Genentech, Inc.
    STREET: 460 Point San Bruno Blvd
    CITY: South San Francisco
    STATE: California
    COUNTRY: USA
    ZIP: 94080
  COMPUTER READABLE FORM:
    MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: patin (Genentech)
  CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,815
  FILING DATE: 15-MAY-1995
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/330128
  FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
  NAME: Torchia, Timothy E.
  REGISTRATION NUMBER: 36,700
  REFERENCE/DOCKET NUMBER: 920C1
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-8674
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 4165 bases
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
US-08-440-815-1

Db   1453 CAA 1455

RESULT 13
US-08-673-789-1
; Sequence 1, Application US/0873789
; Patent No. 5814479
; GENERAL INFORMATION:
; APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
; APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
; APPLICANT: GEORGE, F.
; TITLE OF INVENTION: BSK RECEPTOR LIKE
; TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,789
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/177,812
; FILING DATE: 04-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4322
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
US-08-673-789-1

Query Match   48.1%;   Score 317.4; DB 2;   Length 4322;
Best Local Similarity 68.5%;   Pred. No. 8.1e-95;
Matches 454; Conservative   0; Mismatches 206; Indels   3; Gaps   1;

```
Qy        1 tgggaggaagatcagtggtgtgggatgaacattacacacccatccagaacttaccagtgtgc  60
            |||||||||||||||| ||||| |||||| ||| ||||||||||||| |||||||| |||||
Db      670 TGGGAAGAGATTGGTGAAGTTGATGAGAACTTATGCCCCATCCACACATACCAAGTGTGC  729

Qy       61 aatgtcatggcccacagtcaaacaattgctgagaacaaactgggtccccaggggaactca  120
            || |||||||| ||||||||| |||||||||||| ||||  || ||| ||||| ||||||
Db      730 AAAGTTATGAACAGAANATCAGAATAANTTGGCTGTGCTGTTGGATNCTCTNAACGAAGGT  789

Qy      121 gctcagaagattatgtggagctcaagttcaagctgcaatagcaattccattg  180
            ||||| || ||||||| || |||||||||||||||||||||||||||||||
Db      790 GCTTCCAGAANTCTTTATTGAACTCAAGTTTACTTTPAAGGGACTCGCAACAGCCTTCCTGGA  849

Qy      181 gtttaggaacttgcaagggagacattcaaacctgtactacactggattggatgtctgatcat  240
            || | ||||||| |||||||||||| ||| ||| ||||||| | |||| || |||||||
Db      850 GGACTCGGGGACTTGTAAGGAGACATTAACAtGTAtTTAttTTTGAATCAGATGATGAGAAT  909

Qy      241 ggggtgaaattcgagagcatcagttacaaagattgacaccattgcagctgatgaaagt  300
            || | |||||||||||||| |||||||||||||  |||||||| ||||||||||||||
Db 1330 CAA 1332
```

RESULT 15
US-08-702-367A-12
Sequence 12, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
CLASSIFICATION:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
US-08-702-367A-12

Query Match   45.9%;   Score 303; DB 4;   Length 3162;
Best Local Similarity   67.1%;   Pred. No. 3.8e-90;
Matches   445; Conservative   0;   Mismatches 215;   Indels   3;   Gaps   1;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2000, 12:08:32 ; Search time 111.14 Seconds
                 (without alignments)
                 1485.755 Million cell updates/sec

Title: US-09-104-340-8
Perfect score: 660
Sequence: 1 tgggaagagatacaagtggtgt...............aaagaggttttatgtgcaa 660

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	3132	1	Q34513	HEK coding sequenc
2	463.2	70.2	3254	1	Q90659	Eph-related PTK Ce
3	327	49.5	3116	1	T02948	EPH-like receptor
4	327	49.5	3348	1	T03100	Protein tyrosine-k
5	321.2	48.7	3592	1	V70208	Rat receptor tyros
6	317.4	48.1	3906	1	V70207	Rat receptor tyros
7	317.4	48.1	4165	1	T18893	Rat REK7 cDNA. AL-
8	317.4	48.1	4322	1	V58192	Mouse Bsk receptor
9	303	45.9	3162	1	T02947	EPH-like receptor
10	303	45.9	4529	1	T02949	EPH-like receptor
11	298.2	45.2	2323	1	T32962	Mouse developmenta
12	298.2	45.2	2901	1	T32961	Mouse developmenta
13	298.2	45.2	4304	1	T32960	Mouse developmenta
14	245.2	37.2	3056	1	Q90662	Eph-related PTK Ce
15	245.2	37.2	3059	1	Q90653	Eph-related PTK Ce
16	245.2	37.2	3125	1	Q90661	Eph-related PTK Ce
17	219.6	33.3	2820	1	Q90654	Eph-related PTK Ce
18	202.2	30.6	4281	1	T53471	elk cDNA. Expressi
19	192	29.1	3105	1	T07308	Receptor tyrosine
20	192	29.1	3105	1	T84528	Mouse Nuk tyrosine
21	186.8	28.3	4049	1	Q90660	Eph-related PTK Ce
22	186.8	28.3	4097	1	Q90657	Eph-related PTK Ce
23	168	25.5	2962	1	T02946	EPH-like receptor
24	138.4	21.0	3546	1	Q90656	Eph-related PTK Ce
25	138.4	21.0	3591	1	Q90655	Eph-related PTK Ce
26	119.2	18.2	3776	1	Q90655	Eph-related PTK Ce
27	119.2	18.1	3751	1	062461	Human embryonal ki
28	117.6	17.8	2982	1	Q90971	Protein p140 cDNA
29	117.6	17.8	4027	1	Q90972	Protein p140 cDNA
30	117.6	17.8	4027	1	Q90982	Protein p140 cDNA
31	115.2	17.5	3663	1	T72320	Embryonic stem cel
32	98	14.8	3133	1	Q90652	Eph-related tyrosi
33	80.6	12.2	3969	1	Q49757	pTK gene HpTK5. Ne
34	79.8	12.1	3969	1	T03099	Protein tyrosine-k
35	79.8	12.1	4290	1	Q92641	Human non-differen
36	79.8	12.1	4290	1	T18394	Receptor type tyro
37	79.8	12.1	4290	1	T42593	Coding sequence fo
38	79.8	12.1	4290	1	T51235	Receptor-type tyro
39	79.8	12.1	450	1	V90061	EST clone CW960. N
40	48.8	7.4	3673	1	V33697	Human thymus recep
41	48.8	7.4	4022	1	V62177	Human receptor typ
42	37.4	5.7	324	1	Q61342	Human brain Expres
43	31.4	4.8	4822	1	V04445	Sequence used in d
44	31.2	4.7	786	1	V43665	Non-inducible Immu
45	31.2	4.7	786	1	V46279	A. thaliana NIM-1

ALIGNMENTS

RESULT 1
Q34513

ID   Q34513 standard; DNA; 3132 BP.
AC   Q34513;
DT   24-MAY-1993 (first entry)
DE   HEK coding sequence.
KW   Primer; expression vector; extracellular domain; human; HEK;
KW   eph/elk-like; kinase; pre-B; cell; T; tumour; lymphoid; LK63;
KW   Lila-1; JM; epithelial; HeLa; receptor-type; thymidine kinase;
KW   TK; ligand; B;cellular response; growth; differentiation; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   5'utr   1..99
FT   /*tag= a
FT   cds   100..3051
FT   /*tag= b
FT   3'utr   3052..3132
FT   /*tag= c
FT   signal_peptide 100..159
FT   /*tag= g
FT   /note= "Purified HEK protein #2"
FT   misc_feature   1723..1795
FT   /*tag= d
FT   mat_peptide   160..216
FT   /*tag= e
FT   /note= "Potential transmembrane region"
FT   mat_peptide   2617..2679
FT   /*tag= f
FT   /note= "Purified HEK protein #1"
PN   WO9300425-A.
PD   07-JAN-1993.
PF   19-JUN-1992; AU0294.
PR   21-JUN-1991; AU-006841.
PR   12-DEC-1991; AU-009992.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Boyd AD, Simpson R, Ward LD, Wicks I, Wilkinson D;
DR   WPI; 93-036373/04.
DR   P-PSDB; R31466.
PT   Receptor-type tyrosine kinase reactive with monoclonal antibody
PT   III-A4 - is EPH-ELK-like kinase, useful for phosphorylating
PT   proteins in modulating pre-B, B and T cell function, in cancer
PT   therapy etc.
PS   Claim 6; Fig 1; 58pp; English.
CC   This sequence encodes human eph/elk-like kinase (HEK). HEK is
CC   expressed in both pre-B cells and T cell lines and in a number of
CC   tumours of human origin, eg. lymphoid tumours LK63, Lila-1 and JM,
CC   and the epithelial tumour HeLa. This receptor-type thymidine kinase
CC   (TK) and/or its ligands are useful as agents in modulation of the
CC   production and/or function of pre-B, B and T cells. The TK and its
CC   analogues have activity in transducing signals or in stimulating
CC   cellular responses such as growth and/or differentiation.
SQ   Sequence 3132 BP;

Query Match   100.0%; Score 660; DB 1; Length 3132;
Best Local Similarity   100.0%; Pred. No. 1.4e-209;
Matches 660;   Conservative   0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
Q90659

```
Qy     1  tgggaagagatcaagtggtggatgaacattacacacaccaccaccatcaaggacttacaggtgtgc   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253  TGGGAAGAGATCAGTGGTGTGTGGATGAACATTACACACACCACCACTTACCAGGTGTGC       312

Qy    61  aatgtcatgaccacagtcaaaacaattggctgagaacaaactggtcccagagaactca      120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313  AATGTCATGGACCACAGTCAAAACAATTGGCTGAGAACAAACTGGGTCCCAGGAACTCA     372

Qy   121  gctcagaagatttatgtggagctcaagttcactctacgagactgcaatagcattccatg    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373  GCTCAGAAGATTTATGTGGAGCTCAAGTTCACTCTACGAGACTGCAATAGCATTCCATG    432

Qy   181  gtttaggaacttgcaaggagaagacattcaacctgtactacataggtggtgatgatcat    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433  GTTTTAGGAACTTGCAAGGAGAACATTCAAACCTGTACTACATAGGTGGTGATGATCAT     492

Qy   241  gggtgaaattcgagagcatcaagtttacagacacccattgcagctcagctgatcattg   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493  GGGGTGAAATTTCGAGAGCATCAGTTTACAACACATTGCAGCTCAGTTGATCATTG       552

Qy   301  ttcaactccaaatggatctctgggacccgtattctgaagctcaaacactgagattaga    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553  TTCACTCAAATGGATCTTGGGACCCGTATTCTGAAGCTCAAACACTGAGAGTTAGAGAAGTA  612

Qy   361  ggtcctgtcaacaagaaggggatttattctgaagactgcagctctggtgttctgttgcc    420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   613  GGTCCTGTCAACAAGAAGGGGATTTATTTGGGCATTTACAAGACTTGCTTGTGTGTGCC    672

Qy   421  ttggtgtctgtgagatatactttcaagaaaagtgcccattttacagtgaagaatctgctatg  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673  TTGGTGTCTGTGTGAGTATACTTTCAAGAAAAGTGCCCATTTTACACTGAGAAATCTGGCTATG  732

Qy   481  tttccagacacggtaccccagtcccctggtgtgaggttagagggtcttgtgtc      540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   733  TTTTCCAGACACGGTACCCCAGTCCCCTGGTTGGGAGGTTAGAGGGTCTTGTGTC      792

Qy   541  aacaaatcttaaggaggagaatcctccaagagtgagtgtactgctggctt      600
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   793  AACAAATCTTAAGGAGGAGAAGATCCTCCAAGGATGTAGTGTACTGGGCTT     852

Qy   601  gtaccattggcaagtgttcctgcaatgctgaagaaagaaggtttatgtgccaa      660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   853  GTACCCATTGGCAAGTGTTCCTGCTATGATGAAGAAGAGGTTTTATGTGCCAA      912
```

CC   embryonic tissues, as well as in the adult brain and retina.
SQ   Sequence 3254 BP; 926 A; 737 C; 796 G; 795 T;

Query Match   70.2%; Score 463.2; DB 1; Length 3254;
Best Local Similarity   81.4%;
Matches 537; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy     1  tgggaagagatcaagtggtggatgaacattacacacaccaccaccatcaaggacttacaggtgtgc   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182  TGGGAAGAGATTAGTGGGTGTTGATGAGGCAGATTATACTCCAATCAGAACTTACCAAGAGAC  241

Qy    61  aatgtcatgaccacagtcaaaacaattggctgagaacaaactggtcccagagaactca      120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242  AATGTTATGGATCACAGTCAAAACAATTGGCTGCGAACAAACTGGATCCAGGAACTCA     301

Qy   121  gctcagaagatttatgtggagctcaagttcactctacgagactgcaatagcattccatg    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302  GCCCAGAAGATTTATGTGGAGCTCAAGTTTACCTTGAGGGACTGCAATAGTATCCCCTA   361

Qy   181  gtttaggaacttgcaaggagaagacattcaacctgtactacataggtggtgatgatcat    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362  GTTCTGGGCACTTGCAAAGAGAGACTTTCAATCTGTATTACATGGAATTCGGATGGACAT  421

Qy   241  gggtgaaattcgagagcatcaagtttacagacacccattgcagctcagctgatgaagt   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422  TTGGCAAAGTTCAGAGAGCACCAATTACAAGAAGATTGACACCATGGGCGGGCTGAGAGAGC  481

Qy   301  ttcaactccaaatggatctctgggacccgtattctgaagctcaaacactgagattaga    360
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   482  TTCACCCAGATGATCTTGGGGACCAGATTCTCAAAGCTGAATACCGAGGTG        541

Qy   361  ggtcctgtcaacaagaaggggatttattctgaagactgcagctctggtgttctgttgcc    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542  GGACCTGTTGAGTAAGGAGCTTTTACTTGGCTTTCAAGAGTGTCATGTGTTGCC        601

Qy   421  ttggtgtctgtgagatatactttcaagaaaagtgcccattttacagtgaagaatctgctatg  480
          |||||||||||||||||||||||||||||||||||||||||||||
Db   602  TTAGTCTCGGTGCGAGTGTACTTCAAGAAGAGTGCCCTTTCAACGTCAAGAAACCTCGCCATG  661

Qy   481  tttccagacacggtaccccagtcccctggtgtgaggttagagggtcttgtgtc      540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662  TTTCCAGATACAGTTCCTATGGACTCCCAGTCCCCTGGGTGGGAGGTTGCGGGGTCTTCTTGTGTC  721

Qy   541  aacaaatcttaaggaggagaatcctccaagagtgagtgtactgctggctt      600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   722  AATCATTCCAAGGAGGAAGAGCCCAAGAGTGTACTCGCAGCACGGAAGTGGGCTA        781

Qy   601  gtaccattggcaagtgttcctgcaatgctgaagaaagaaggtttatgtgccaa      660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   782  GTGCCCATAGGGAAGTGCTTGTGTGTAATGTCTGGCTATGAAGAGAGAGGAGGCTTTGCGTGCCAA  841
```

RESULT 3
T02948
ID   T02948 standard; cDNA; 3116 BP.
AC   T02948;
DT   16-APR-1996 (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK8 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK8;
KW   human eph-like receptor protein kinase; therapy; diagnosis; antibody; vector; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds   34..2994
FT   /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PD   14-APR-1995; U04681.
PF   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PR   Fox GM; Jing S, Welcher AA;
PI   WPI; 95-373799/48.
DR   P-PSDB; R85091.

ID   Q90659 standard; DNA; 3254 BP.
AC   Q90659;
DT   11-NOV-1995 (first entry)
DT   07-SEP-1994; U10140.
DT   03-DEC-1993; US-162809.
DE   Eph-related PTK Cek4.
KW   Cek4; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
OS   Gallus sp.
FH   Key   Location/Qualifiers
FT   cds   32..2980
FT   /*tag= a
PN   WO9515375-A.
PD   08-JUN-1995.
PS   Disclosure; Page 85-89; 129pp; English.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
PR   Pasquale EB, Sajjadi FG;
PI   WPI; 95-215256/28.
DR   P-PSDB; R75711.
CC   Probes derived from the EPH-related PTKs Cek4 (Q90659) and Cek5
CC   (Q90660) were used to isolate novel cDNA clones (Q90652-58,
CC   Q90661-62) from chicken embryo and embryonic brain libraries.
CC   Cek4 is highly expressed in the chicken developing brain and

PT New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT and related vectors, host cells, proteins, antibodies etc., used
PT diagnostically and therapeutically to modulate receptor activation
PS or prodn.
PT Claim 1; Page 57-62; 133pp; English.
CC New human EPH-like receptor protein tyrosine kinase
CC cDNAs (TO2946-49) coding for 4 novel human EPH-like receptor protein
CC tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC were isolated from a human foetal brain cDNA library using a directed
CC PCR approach with primers (see TO2960-61) based on conserved regions of
CC extensive homology to the EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC homology to the catalytic domain of chicken EPH-like
CC receptor PTKs and EPH-like receptor, in hybridisation assays, and
CC EPH-like receptor. The isolated cDNAs are used for prodn. of
CC receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC to detect abnormalities in HEK receptor genes.
SQ Sequence 3116 BP; 859 A; 720 C; 812 G; 725 T;

Query Match 49.5%; Score 327; DB 1; Length 3116;
Best Local Similarity 69.4%;
Matches 460; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy    1 tgggaagaagatcaagtggtgtgatgaacattacacaccatcacaggacttacaggtgtgc  60
         |||| |  | |||||||||||||||||||||  ||| |||| ||  |||||||||||||
Db  193 TGGGAGGAAGTGAGTRNCATCGGATRCATGGATGAAAAAAATACACCAATCCGAACCTACCAAGTGTGC 252

Qy   61 aatgtcatggaccacagtcaaaacaaattggctgagaacaaactgggtccccaggaactca 120
         ||  |||||||  || || |  |  ||| ||| ||| |||||||  ||||||||||||||
Db  253 AANTGTCATGGAACCCACAGCCAGAATAACTGGGCTACGAACTGANTGGANTCACCCGAGAAGGG 312

Qy  121 gctcagaagattatgtggagctcaagttcactacgagacatgcaatgcaattccatg 180
         |||| |||| ||| ||| || |||||| ||||| |||| ||| || |  |||||||
Db  313 GCTCAGAGGGTGTATATTGAGATTAAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC 372

Qy  181 gtttaggaacttgcaaggagaccattcaaacctgtactacatggagtctgatgatcat 240
         ||  |||||||  || | |||||||||||||||||||||||||||| ||||||||||
Db  373 GTCATGGGGACTTGCAAGGACGGTTTAAACCTGTACTACTATGAAATCAGACAACGACAAA 432

Qy  241 ggggtgaaattcgagagcatcagttacaaagattgaccagctcgatgatgaaagt 300
         |||||||||||||||||||  |||||||||  || ||||||||  |||||||||
Db  433 GAGCGGTTCATCAGAGNACACAGTTTGTCAAAATTGACACACATTGCTGCTGATGAGAGC 492

Qy  301 ttcactcaaatggatcttggggacgtattctgaagctcaacactgattagagaagta 360
         ||||||||||||||||| | ||||| ||||| ||||||  ||||| ||| ||||||
Db  493 TTCACCCAAGTGGACANTGGTGACAATCATGAAGCTGAAACACCAACCAATCAGAACAAAGTGTCA 552

Qy  361 ggtccgtcaacaagaaggggattatatttggcattcaaagatgttggtgttgtgcc 420
         |||| ||||||||||||  ||| ||| || ||||||| || ||||| | |||||
Db  553 GGGGCAATAAGCACAAAAAGGGGTTTTACCTGCGTTTCAGGANTGTGGGGGCTCCATCGCC 612

Qy  421 ttggtgtctgtgagagtatacttcaaaaaagtgccccatcattacagtggctatg 480
         ||||| ||| ||| ||||| ||  |||||| | ||||||||||||||||||||
Db  613 CTGGTATCAGTCCGTGTGTTCTATAAAAAGTGTCCACTCACAGACTGGCCGG 672

Qy  481 tttccagacacggtacccatg---gactccagtccctgggaggggttcttgt 537
         |||||||||||||||  ||||     |||  |||||||||||||||| ||||
Db  673 TTTCCTGACCATCACAGGAGCATGGGGTGATCCGGGATCCGGGATGTA 732

Qy  538 gtcaacaatctaaggaggaagagatctccaaggagatgtactcagcgaatgg 597
         || || ||||| ||||||||||| |||| ||||| ||| |||||||| ||
Db  733 GTCACAGAACTCAGAGAGAGAAAAGATGTGTCCAAAAATGTACTGTGGGGGCGAATGG 792

Qy  598 cttgtcaccattggcaagtgtcctgcatgatgagaaagagtttatgtgc 657
         ||||| ||||||| ||||| || ||||||||||||| ||||| || ||
Db  793 CTGGTACCCATTGGCAACTGCTATCCAACCCTTACCAGAGGAGGGGAGAAATGC 852

Qy  658 caa 660
         |||
Db  853 CAA 855
```

ID T03100 standard; DNA; 3348 BP.
AC T03100;
DT 14-FEB-1996 (first entry)
DE Protein tyrosine-kinase bpTK7 gene.
KW Protein tyrosine-kinase; pTK; bpTK7; agonist; cell growth;
KW differentiation; ss.
OS Homo sapiens.
FH Key   Location/Qualifiers
FH
FT cds   1..2961
FT   /*tag= a
FT signal_peptide   1..57
FT   /*tag= b
FT mat_peptide   58..2958
FT   /*tag= c
CC Protein tyrosine kinase (pTK) genes.
PN WO9527061-A1.
PD 12-OCT-1995; US; U04228.
PD 04-APR-1995; US-222616.
PR 04-APR-1994; US-222616.
PA (GETH) GENENTECH INC.
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI Wood WI;
DR WPI; 95-366160/47.
DR P-PSDB; R85936.
PT Agonist antibodies which activate specific protein tyrosine
PT kinases(s) - also activate chimeric proteins of kinase extracellular
PT domain and Ig constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
PS Disclosure; Page 88-92; 125pp; English.
CC DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC to screen cDNA libraries to identify novel pTK genes. The bpTKs,
CC bpTK1, bpTK2, bpTK3, bpTK4, bpTK5 and R85924-28 and R85935,
CC respectively) are expressed in human brain tissue and show homology
CC to known pTKs. A full-length sequence for the bpTK7 gene (T03100)
CC was obtd. This gene may be used to design new drugs, peptides and
CC antisense constructs that modulate pTK activity.
SQ Sequence 3348 BP; 924 A; 855 G; 800 T;

Query Match 49.5%; Score 327; DB 1; Length 3348;
Best Local Similarity 69.4%;
Matches 460; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy    1 tgggaagaagatcaagtggtgtgatgaacattacacaccatcacaggacttacaggtgtgc  60
         |||| |  | |||||||||||||||||||||  ||| |||| ||  |||||||||||||
Db  160 TGGGAGGAAGTGAGTRNCATCGGATRCATGGATGAAAAAAATACACCAATCCGAACCTACCAAGTGTCC 219

Qy   61 aatgtcatggaccacagtcaaaacaaattggctgagaacaaactgggtccccaggaactca 120
         ||  |||||||  || || |  |  ||| ||| ||| |||||||  ||||||||||||||
Db  220 AANTGTGATGGAACCCACAGCCAGAATAACTGGGCTACGAACTGANTGGANTCACCCGAGAAGGG 279

Qy  121 gctcagaagattatgtggagctcaagttcactacgagacatgcaatgcaattccatg 180
         |||| |||| ||| ||| || |||||| ||||| |||| ||| || |  |||||||
Db  280 GCTCAGAGGGTGTATATTGAGATTAAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC 339

Qy  181 gtttaggaacttgcaaggagaccattcaaacctgtactacatggagtctgatgatcat 240
         ||  |||||||  || | |||||||||||||||||||||||||||| ||||||||||
Db  340 GTCATGGGGACTTGCAAGGACGGTTTAAACCTGTACTACTATGAAATCAGACAACGACAAA 399

Qy  241 ggggtgaaattcgagagcatcagttacaaagattgaccagctcgatgatgaaagt 300
         |||||||||||||||||||  |||||||||  || ||||||||  |||||||||
Db  400 GAGCGGTTCATCAGAGNACACAGTTTGTCAAAATTGACACACATTGCTGCTGATGAGAGC 459

Qy  301 ttcactcaaatggatcttggggacgtattctgaagctcaacactgattagagaagta 360
         ||||||||||||||||| | ||||| ||||| ||||||  ||||| ||| ||||||
Db  460 TTCACCCAAGTGGACANTGGTGACAATCATGAAGCTGAAACACCAACCAATCAGAACAAAGTGTCA 519

Qy  361 ggtccgtcaacaagaaggggattatatttggcattcaaagatgttggtgttgtgcc 420
         |||| ||||||||||||  ||| ||| || ||||||| || ||||| | |||||
Db  520 GGGGCAATAAGCACAAAAAGGGGTTTTACCTGCGTTTCAGGANTGTGGGGGCTCCATCGCC 579

Qy  421 ttggtgtctgtgagagtatacttcaaaaaagtgccccatcattacagtggctatg 480
         ||||| ||| ||| ||||| ||  |||||| | ||||||||||||||||||||
Db  580 CTGGTATCAGTCCGTGTGTTCTATAAAAAGTGTCCACTCACAGACTGGCCGG 639
```

RESULT 4
T03100

```
Qy   481  tttcccagacacggtacccatg---gactcccagtccctggtggaggttagagaggtcttgt 537
          ||||||    ||||||||||||||   |||||||||||||||||||||||| ||||||||||||
Db   640  TTTCCCGACACCATTACTAGAGGGGCCTGATAGCTCTTCCCTGGTGGAAGTTCGAGGCTCCTGT 699

Qy   538  gtcaacaaattctaagagaggaggaagatctccaaggatgtactgcagtacagagaggcgaatgg 597
          ||||||| || ||||||||||||||| || ||||||||||||||||||| ||||||||||||||
Db   700  GTCAACAACTCAGAAGAGAGGAAGAAATGTGCCAAAAATGTACTGTGTGGGCAGATGGTGAATGG 759

Qy   598  cttgtaccattggcaagtgttcctgcaatgctgtatgtaagagaggagggttttatgtgc 657
          |||||||||||||||||||| || ||||||||||| || ||||||||||||||||||||||
Db   760  CTTGGTACCCATTGGCAACTGCCTATGCAACCGTGGGACGGCGGAGCGGGAGGGAAATGC 819

Qy   658  caa 660
          |||
Db   820  CAA 822
```

```
RESULT 5

ID   V70208    standard; DNA; 3592 BP.
AC   V70208;
DT   11-FEB-1999  (first entry)
DE   Rat receptor tyrosine kinase Ehk-2 encoding DNA.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis; ss.
OS   Rattus sp.
FH   Key            Location/Qualifiers
FT   CDS            598..3444
FT                  /*tag= a
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995; US-469537.
PR   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Masiakowski P, Yancopoulos GD;
WPI  99-044584/04.
P-PSDB; W83148.
DR   
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence encodes rat Ehk-2.
SQ   Sequence 3592 BP;    946 A;    916 C;    894 G;    834 T;

PN   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Claim 7; Fig 21; 194pp; English.
DR   P-PSDB; W83148.

Query Match       48.7%;   Score 321.2;   DB 1;   Length 3592;
Best Local Similarity     68.8%;
Matches 456;   Conservative   0;   Mismatches 203;   Indels   3;   Gaps   1;

Qy     1  tgggaggagatcagtgtgtgatgaacatacacaccacaggactacaaggtgtgc 60
          ||||| || ||| ||||| ||||||||| || |||||| || |||||||||||||
Db   763  TGGGATGCCATTACTGAAATGGATGAACACAGGCCCATACACATACCAGGTATGC 822

Qy    61  aatgtcatgaccacagtcaaacaattggctgagaacaaactgggtcccccaggaactca 120
          ||||| ||||| || ||||| || |||||||| ||||| |||| |||||||||||||
Db   823  AATGTCATGGGACCAAACACAACTGGCTCGTACTAACTGGGATCTCGTGATGCT 882

Qy   121  gctcagaagaatttatgtgggggctcaagttcactctacgagagactgcaatagacattccattg 180
          || |||||||| || |||||| ||||||||||| ||||| |||||| |||| ||| ||||| |
Db   883  GCTCAGAAAATCTATGTGGGAAATGAAGTTCACATTGAGAGATTGTAACAGACATCCCATGG 942

Qy   181  gttttaggaacttgcaaggagacacattcaacctgtactacacatggagtctgatgtgatcat 240
          || ||||| ||||| ||||| || || |||||||||| |||||| || ||||||| ||||||
Db   943  GTCTTGGGAACGTGTAAAGAAACATTTACTCTGTATATTACAGAAATCGACGAATCCCAC 1002

Qy   241  ggggtgaaattccgagagcatcagttacaagaattgacaccattgcagctgatgaaagt 300
```

```
Db  1003  GGAACTAAAATTCAAGCACAGCCAAGCCAATATAAAGAAGATTGACACAATTGCTCGGAGAGACT 1062
          || ||||||||||| || ||||||||| || |||||||| |||||| || |||||| ||||

Qy   301  ttcacctcaaatgaatgctctgggggaccgtatttctgaagctcaaacactgagattag... 360
          |||| ||||| ||||||| ||| |||||| ||| |||||||||| || |||||| ||||
Db  1063  TTTACTCAGATGGAATTGGGTGATCGCCATACCTTAAACTCAACACTGAAGTTCGTGAGGTG 1122

Qy   361  ggtcctgtcaaacaagaaggggatttttattttgggtcattcaagatcgttcttgtgttgcc 420
          |||||| |||| |||||||| |||||||| |||||| || |||||||||||| || |||
Db  1123  GGGCCAATGAACAGGGAAAAGGGGATTCTATTTGGGGATTATTGGACGAGCATCGATGGCT 1182

Qy   421  ttggtgtctgttgagtatactccaaaaagtgcccatttacagtgaagaatctggtcatg 480
          || || ||| || ||||| ||||| ||||||| |||||||| || |||||| ||||| 
Db  1183  CTGGTCTCAGTCCGTGTGTTCTACAGAAAAATGCCCCTTCCACTGTGCCGAGAACTTGGCTANG 1242

Qy   481  tttcccagacacggtaccca---tggactcccagtcccctgggttgaggttagagggtcttgt 537
          ||||||    |||||||||||    |||||||||||| ||||||| ||||||| ||||||||||
Db  1243  TTTCCGGATACCATTCCAAGGGGTTGACTCCTCCCTGGTGGTTGAAGTGCGGGGGGCTCATGC 1302

Qy   538  gtcaacaaattctaagagaggaggaagatccctccaaggatgtactgcagtacagagaggcgaatgg 597
          |||||||| ||| ||||||||| || ||||| |||||||| ||||| ||||| ||||||||||
Db  1303  GTGAAGAGTTCTGAGGACGGAGATACTCCTAAACTCAACTACACTGTGGGAGCGCGATGGGACTGG 1362

Qy   598  cttgtaccattggcaagtgttcctgcaatgctgtatgtaagagaggagggttttatgtgc 657
          |||||||||||||||||||| ||||| ||| || || ||||||||||||||||||||||||
Db  1363  CTCCTCCCTTGGTGAAGGTGTATGTGCACTACCAGGGTATGAAGAAATGCAGGGGGTTCTTGC 1422

Qy   658  ca 659
          ||
Db  1423  CA 1424
```

```
RESULT 6

ID   V70207    standard; DNA; 3906 BP.
AC   V70207;
DT   11-FEB-1999  (first entry)
DE   Rat receptor tyrosine kinase Ehk-1 encoding DNA.
KW   Receptor tyrosine kinase; Ror-1; Ror-2; Ehk-1; Ehk-2; detection;
KW   neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW   binding protein; BDNF; NT-3; diagnosis; ss.
OS   Rattus sp.
FH   Key            Location/Qualifiers
FT   CDS            476..3493
FT                  /*tag= a
PN   US5843749-A.
PD   01-DEC-1998.
PF   06-JUN-1995; US-469537.
PF   17-MAR-1995; US-406247.
PR   26-JUL-1991; US-736559.
PR   28-OCT-1993; US-144992.
PR   06-JUN-1995; US-469537.
PA   (REGE-) REGENERON PHARM INC.
PI   Maisonpierre PC, Masiakowski P, Yancopoulos GD;
WPI  99-044584/04.
P-PSDB; W83147.
DR   
CC   The present invention describes nucleic acid molecules for ror-1,
CC   ror-2, ehk-1 and ehk-2. Also described are the corresponding proteins:
CC   Ror-1; Ror-2; Ehk-1; and Ehk-2. The proteins are orphan receptor
CC   tyrosine kinases. The present sequence encodes rat Ehk-1.
SQ   Sequence 3906 BP;   1097 A;    876 C;    990 G;    943 T;

PN   DNA encoding receptor tyrosine kinase proteins - and corresponding
PT   proteins
PS   Claim 5; Fig 22; 194pp; English.
DR   P-PSDB; W83147.

Query Match       48.1%;   Score 317.4;   DB 1;   Length 3906;
Best Local Similarity     68.5%;
Matches 454;   Conservative   0;   Mismatches 206;   Indels   3;   Gaps   1;

Qy     1  tgggaggagatcagtgtgtgatgaacatacacaccacccatcaaggactacaaggtgtgc 60
          ||||| |||||| ||||| ||||||||| || |||||||||||| |||||||||||||||
Db   728  TGGGAAGAGAATTGGTGAAGTTGATGAGAAAAACTATGCCCCATCCATCACCACTATCAAGGTGTGC 787
```

RESULT 7

```
ID    T18893 standard; cDNA; 4165 BP.
AC    T18893;
DT    05-JAN-1997  (first entry)
DE    Rat REK7 cDNA.
KW    REK7; eph-related tyrosine kinase receptor; AL-1; ligand;
KW    neurotrophic factor; neuropathy; angiogenesis; therapy; diagnosis;
OS    Rattus sp.
FH    Key             Location/Qualifiers
FT    cds             541..3327
FT                    /*tag= a
FT    signal_peptide  541..711
FT                    /*tag= b
FT    mat_peptide     712..3324
FT                    /*tag= c
KW    WO9613518-A1.
PN    09-MAY-1996.
PD    26-OCT-1995;  U14016.
PR    27-OCT-1994;  US-330128.
PR    07-JUN-1995;  US-486449.
PA    (GETH )  GENENTECH INC.
PI    Caras IW,  Winslow JW;
DR    WPI;  96-239448/24.
DR    P-PSDB;  R97853.
PT    AL-1 neurotrophic factor, an eph-tyrosine kinase ligand - used in
```

Left-column alignment:

```
Qy   61 aatgtcatggaccacagtcaaacaatttggctgagaacaacaaactgggtcccagggaactca 120
        ||||| |||||| ||  ||||| || ||||   |||||    |||||  ||||||||||||
Db  788 AAAGTTATTGGAACAGAATCAGAATAATTGGCTGTGTTGGATCCTGATTCTCTAACGAAGGT 847

Qy  121 gctcaagaagatttatgtggagctcaagttcaatagcatagacattccattg 180
Db  848 GCTCCAGAATTTTATTGAACTCAAGTTACTCTGAGGGATTGCAACAGCCTTCCTGGA 907

Qy  181 gttttaggaacttgcaagaggagacattcaacctgtactacatgatgctgatgatcat 240
Db  908 GGACTGGGGACTTGCAAGGAGGACCTTAACATGTATATTATTTGAGTCGGATGAGAAT 967

Qy  241 ggggtgaaattcgagagcatcagttacaaagattgacaccattgcagctgatgaaagt 300
Db  968 GGGAGAAAATATCAAAGAGAACCAGTACATCAAGAAGATACCAATTGCTCGTGATGAGAGC 1027

Qy  301 ttcactcaaatggatcttgggaggaccgtattctgaagctcaacactgagattagagaagta 360
Db 1028 TTCACGGAACTTGACCTTGGACACCGGGTCATGAAGCTGAATACGGAGGGTCAGAGATGTA 1087

Qy  361 ggtcctgtcaacagaaggggdattttatttggcattccaagtagtgttggtcttgttgcc 420
Db 1088 GGACCTCTGAGCAAAAAGGGATTTTANCTTCGCTTCCAAGAGTGCCATCACCTATCAAAGTGTGC 1147

Qy  421 ttggtgtctgtgagagtatacttcaaaaagtgcccattgaagaatctctggctatg 480
Db 1148 CTGGTTTCTGTCCGTGTCTACTATAAAAAAGTGTCTTCTCTAGTTGAACATTTGGCTGTGTT 1207

Qy  481 tttccagacacggtacccatggactcccagtccc--tggtggaggttagaggtcttgt 537
Db 1208 TTCCCTGACACGATCACTGCAGATTCTTCCAGTTGCTAGAGGGATTGCAACAGCGCTTCCIGC 1267

Qy  538 gtcaacaattcctaaggagaggaagatcctccaagagtgtactgcgaaggcgaatg 597
Db 1268 GTCAACCAATTCTGTACAGACGATCCTCCCAAAAATCCATCAGTGCTTGCAAGTGGGATGAGAAT 1327

Qy  598 cttgtgtacccattggcaagtgttcctgcaatgctggctatgaagaaagagagttttatgtgc 657
Db 1328 CTGGTTCCATCGGGAAATCGCCGGAAAAATGGCAGAGAGAAAAATGGTACCTGT 1387

Qy  658 caa 660
        |||
Db 1388 CAA 1390
```

```
SQ    Sequence 4165 BP;   1157 A;   907 C;   1035 G;   1066 T;
CC    brain cDNA library, leading to the isolation of AL-1 cDNA (T18897).
CC    cDNA, and an amplified fragment was used to screen a human foetal
CC    ligands were used to amplify human breast carcinoma BT20 cell
CC    expression of REK-IgG fusion was used to screen cultured cell lines for surface
CC    REK-IgG fusion was used to screen cultured cell lines for isolated
CC    full-length REK7 cDNA from a rat hippocampal cDNA library. An
CC    cDNA library (T18894-96) to amplify cDNAs of an adult mouse hippocampal
CC    primers (T18894-96) to amplify cDNAs of an adult mouse hippocampal
CC    ligand.  It was isolated using degenerate receptor tyrosine kinase
CC    A cDNA clone (T18893) codes for rat REK7 (R97853), an eph-related
CC    Example 1; Page 47-49; 75pp; English.
PS    tyrosine kinase receptor, for which AL-1 (see also W97854) is a
PT    angiogenesis-related conditions.
PT    treatment and diagnosis of neuronal disorders and
```

Middle-column query match:

```
Query Match     48.1%;  Score 317.4;  DB 1;  Length 4165;
Best Local Similarity  68.5%;
Matches  454;  Conservative  0;  Mismatches  206;  Indels  3;  Gaps  1;
```

Right-column alignment:

```
Qy    1 tgggaagaagatcaagtggtgtcggatgaacacatccacacactcaggagacttaccagtgtgtgc 60
Db  793 TGGGAAGAGAGATTGGTGAAGTTGATGAAAACATTCCCCCATCACACCTATCAAAGTGTGC 852

Qy   61 aatgtcatggaccacagtcaaacaatttgggtctgagaacaacaaactgggtcccagggaactca 120
Db  853 AAAGTTATTGGAACAGAATCAGAATAATTGGCTGTGTTGGATCCTGATTCTCTAACGAAGGT 912

Qy  121 gctcaagaagatttatgtggagctcaagttcaatctcacgctacgagactgcaatagacattccattg 180
Db  913 GCTTCAGAAATTTTATTGAACTCAAGTTACTCTGAGGGATTGCAACAGCGCTTCCTGC 972

Qy  181 gttttaggaacttgcaagaggagacattcaacctgtactacatgatgatctgatgtgatcat 240
Db  973 GGACTGGGGACTTGCAAGGAGGACCTTAACATGTATATTATTTGAGTCGGATGAGAAT 1032

Qy  241 ggggtgaaattcgagagcatcagttacaaagattgacaccaattgcagctgatgaaagt 300
Db 1033 GGGAGAAAATATCAAAGAGAACCAGTACATCAAGAAGATACCAATTGCTCGTGATGAGAGC 1092

Qy  301 ttcactcaaatggatcttgggaggaccgtattctgaagctcaacactgagattagagaagta 360
Db 1093 TTCACGGAACTTGACCTTGGACACCGGGTCATGAAGCTGAATACGGAGGGTCAGAGATGTA 1152

Qy  361 ggtcctgtcaacagaaggggdattttatttggcattccaagtagtgttggtcttgttgcc 420
Db 1153 GGACCCTGTGAGCAAAAAGGGATTTTANCTTCGCTTCCAAGAGTGCCATCACCTATCAAAGTGTGC 1212

Qy  421 ttggtgtctgtgagagtatacttcaaaaagtgcccattgaagaatctctggctatg 480
Db 1213 CTGGTTTCTGTCCGTGTCTACTATAAAAAAGTGTCTTCTCTAGTTGAACATTTGGCTGTGTT 1272

Qy  481 tttccagacacggtacccatggactcccagtccc--tggtggaggttagaggtcttgt 537
Db 1273 TTCCCTGACACGATCACTGCAGATTCTTCCAGTTGCTAGAGGGATTGCAACAGCGCTTCCTGC 1332

Qy  538 gtcaacaattcctaaggagaggaagatcctccaagagtgtactgcagaaggcgaatg 597
Db 1333 GTCAACCAATTCTGTACAGACGATCCTCCCAAAAATCCATCAGTGCTTGCAAGTGGGATGAGGTGG 1392

Qy  598 cttgtgtacccattggcaagtgttcctgcaatgctggctatgaagaaagagagttttatgtgc 657
Db 1393 CTGGTTCCATCGGGAAATCGCCGGAAAAATGGCAGAGAGAAAAATGGTACCTGT 1452

Qy  658 caa 660
        |||
Db 1453 CAA 1455
```

RESULT 8

```
ID    V58192 standard; cDNA; 4322 BP.
```

AC V58192;
DT 25-NOV-1998 (first entry)
DE Mouse Bsk receptor-like tyrosine kinase cDNA clone.
KW Mouse; Bsk; receptor-like tyrosine kinase; brain; diagnosis; neoplasm;
KW neurodegenerative disease; limbic system neuron regeneration;
KW chromosomal abnormality; degenerative growth; development disorder;
KW viral infection; bacterial infection; Alzheimer's disease; epilepsy;
KW schizophrenia; stroke; cerebral ischaemia; ds.
OS Mus sp.
FH Key   Location/Qualifiers
FT CDS   418..3051
FT   /*tag= a
FT   /product= "Bsk"
FT   /note= "receptor-like tyrosine kinase"
PN US5814479-A.
PD 29-SEP-1998.
PF 11-JUN-1996; 673789.
PR 04-JAN-1994; US-177812.
PR 11-JUN-1996; US-673789.
PA (KROM/) KROMER L F.
PA (SCHU/) SCHULZ N T.
PA (WOUD/) WOUDE G F V.
PA (ZHOU/) ZHOU R.
PI Kromer LF, Schulz NT, Woude GFV, Zhou R;
DR WPI; 98-541751/46.
DR P-PSDB; W71628.
PT Isolated nucleic acid sequence encoding protein - used in Bsk
PT nucleic acid probes, used in detecting alterations in level of Bsk
PT messenger-RNA in biological samples isolated from mammal afflicted
PT with disease
PS Claim 2; Fig 2; 72pp; English.
CC The present sequence encodes mouse Bsk, which is a receptor-like
CC tyrosine kinase. The nucleic acid sequence encoding Bsk can be used in
CC Bsk nucleic acid probes, which can be used in detecting alterations in
CC the level of Bsk messenger-RNA (mRNA) in biological samples isolated
CC from a mammal afflicted with a disease, such as neurodegenerative
CC diseases or disorders and neoplasms. The nucleic acid sequence can also
CC be delivered into the limbic system of patients with limbic system
CC diseases. The nucleic acid sequence or injury, to promote or enhance
CC limbic system neuron regeneration or growth. Such neurodegenerative
CC diseases include, chromosomal abnormalities, degenerative growth and
CC development disorders, viral infections, bacterial infections, brain
CC injuries, neoplastic conditions, Alzheimer's disease, epilepsy,
CC schizophrenia, or stroke and cerebral ischaemia.
SQ Sequence 4322 BP; 1260 A; 887 C; 1055 G; 1120 T;

Query Match   48.1%; Score 317.4; DB 1; Length 4322;
Best Local Similarity   68.5%; Pred. No. 1; 1.7e-95;
Matches 454; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

```
Qy      1 tgggaagagatcaagtggtgtggatgaacattacacaccatcaggagtta cacaaggtgtgc 60
          ||||||||||||||||||||||||||||||||| || |||||||| |||||||||| || |||
Db    670 TGGGAAGAGATTGGTGAAGTTGATGAGAACTATGCCCCCATCCACACATACCAAGTGTGC 729

Qy     61 aatgtcatggaccacgtcaaaacaattggctgagaacaaactgggtgtcccaggaactca 120
          | ||| ||||| |||||| |||| ||||| || |||||| |||||| ||||||||||||||
Db    730 AAAGTTATGGAACAAGAATCACAAATAATTGGCGTGTTGACCAGTTGGATCTCTACGAAGGT 789

Qy    121 gctcagaagatttatgtgggagctcaagttcactcctacgagactgcaatagccatccattg 180
          |||||||| |||||| |||||||| |||||||||| | |||| |||| |||||||||||| |
Db    790 GCTTCCAGAATCTTTATTGAACTCACGTTTACTTTAAGGGACTGCAACAGCCTTCCTGGA 849

Qy    181 gttttaggaacttgcaaggagagacattccaacctgtactacaaggtctgatgatgatcat 240
          |||| |||||||||||| ||||| || |||||| |||| || ||||||||||||||||| ||
Db    850 GGACTCGGGACTTGTGTAAGGAGACANTTTAACATGTAATATATTTGAATCAGATGATGAGAAT 909

Qy    241 ggggtcgaaatttcagaagcatcagtttacacacaattgcagctgatgaaat 300
          ||| |||||| ||||||||||||||||| |||||||||||||| || ||||||
Db    910 GGGAGAAGTATCAAAAGAGAACAATACCATCAGAGATGATGATACATCGCTGCAGAATGAGAGC 969

Qy    301 ttcactcaaatggatcttgggaccgtattctgaagctcaacactgagagattagagaagta 360
          |||||||||||||||| ||||||||||| ||||||||||||||||||||||||||||||||
Db    970 TTCACAGAAACTTGATCTTGGTGACCGTGTCATGAAGCTCAATACAGAGGTCAGAGAGATGTC 1029

Qy    361 ggtcctgtcaacaacaagagagagatttatttggcattcaagatgttggtcttgttgttgcc 420
          ||||||||||||||||||||||| |||| |||| ||||||||| ||||||| || ||| |||
Db   1030 GGAACCTCGAGCAAAAAAGGATTTTAACTTGCTTTCCAAGATGTCCGGTGCCATTGCT 1089

Qy    421 ttggtgtcctgtgaagtatacttcaaaagtgcccattttacgtggaagaatctggctatg 480
          |||||| ||||| |||| |||||| || | | || ||||| | |||||||| |||||||||
Db   1090 CTGGTTGTTCGTGTCCGTGTCTTGACTATATAAAAAGTGTCCCTCGTACTGTAGTACTGGTATC 1149

Qy    481 tttccagacacgggtacc--catgagactcccgtgtggaggttagaggggttcttgt 537
          ||| |||||| |||||||  ||||| |||||||||||||||||| ||||||||||||||
Db   1150 TTCCCTGACACTATCACTGGACAGAAGATTCATCACAGATTGTTGTGAGGGGTCAGGCTCTGC 1209

Qy    538 gtcaacaattctaaggagaggaagatcctccaaggatgtactgcagttacgaaggcgaatgg 597
          || ||||||| |||| ||||||| |||||||||| ||||||||| | ||||||||||||||||
Db   1210 GTCAACCANTTCTGTGACAGATGATCCTCCCAGATGCATTGACATTGCATTCGCTGGAGAAGGG 1269

Qy    598 cttgtacccatggcaagtgtccctgcaatgtggctggctgggaagaagagagttttatgtgc 657
          ||||| ||| ||||||||||| |||||| || ||||| |||||| || ||| |||| |||||
Db   1270 CTGGTTCCCATTGGGAAATGCATGTGCAAGGCCTGGATATGAGAGAAAAATGGTACCTGC 1329

Qy    658 caa 660
          |||
Db   1330 CAA 1332
```

RESULT 9
ID   T02947 standard; cDNA; 3162 BP.
AC   T02947;
DT   16-APR-1996 (first entry)
DE   EPH-like receptor protein tyrosine kinase HEK7 cDNA.
KW   EPH-like receptor protein tyrosine kinase; PTK; HEK7;
KW   human eph-like kinase; therapy; diagnosis; vector; antibody; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   cds   1..2976
FT   /*tag= a
PN   WO9528484-A1.
PD   26-OCT-1995.
PF   14-APR-1995; U04681.
PR   15-APR-1994; US-229509.
PA   (AMGE-) AMGEN INC.
PI   Fox GM, Jing S, Welcher AA;
DR   WPI; 95-373799/48.
DR   P-PSDB; R85090.
PT   New nucleic acid encoding EPH-like receptor tyrosine kinase(s) -
PT   and related vectors, host cells, proteins, antibodies etc., used
PT   diagnostically and therapeutically to modulate receptor activation
PT   or prodn.
PS   Claim 1; Page 49-54; 133pp; English.
CC   cDNAs (T02946-49) coding for 4 novel human EPH-like receptor protein
CC   tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (R85089-92), respectively,
CC   were isolated from a human foetal brain cDNA library using a directed
CC   PCR approach with primers (see T02960-61) based on conserved regions of
CC   receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC   extensive homology to the catalytic domain of chicken EPH-like
CC   receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC   EPH-like receptor. The isolated cDNAs are used for prodn. of
CC   recombinant HEKs and chimeric receptors, host cells, in hybridisation assays, and
CC   to detect abnormalities in HEK receptor genes
SQ   Sequence 3162 BP; 921 A; 667 C; 775 G; 799 T;

Query Match   45.9%; Score 303; DB 1; Length 3162;
Best Local Similarity   67.1%; Pred. No. 9e-91;
Matches 445; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

```
Qy      1 tgggaagagatcaagtggtgtggatgaacattacacaccatcaggagttacacaaggtgtgc 60
          |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db    175 TGGGAAGAGATTGGTGAAGTGGATGAAGAAATTATGGCCCCTATCCACACACCATACCAAGTATGC 234
```

```
Qy      61 aatgtcaatggaccacagtgaacaaacaattggctgagaacactgggtcccaggaactca 120
           |||||||||||||||||||||||||||||||||||||| |||||| ||  ||||||||||
Db     235 AAAGTGAATGGAACAGAATCAGAATAACTGGCTTTTGACCCAGTTGGAATCCCAATGAAGGT 294

Qy     121 gctcagaagatttatgtggagctcaagttcactctaccacgacactgcaattccatty 180
           ||||||||| |||| ||||| |||||||||||||||||||| |||||||||||||||||
Db     295 GCTTCCAGAAACTTCATAGAACTCAAATTTACCCTGCGGGACTGCAACACCCTTCCTGGA 354

Qy     181 gtttaggaacttgcaaggagacattcaacctgtactacatggagtctgatgatgatcat 240
           ||||||||||||||||||||| |||||||||||||||||||||||||||
Db     355 GGACTGGGGACCTGTAAGGAAACCTTTAATGTGATTACTTTGAGTGTCAGATGAATCAGAAT 414

Qy     241 ggggtgaaatttcagagagcatcagtttacaaaggttgaccacctgcagctgatgaaagt 300
           |||| ||||||||| ||||||||||||| ||||| |||||||||||||||||| ||||||
Db     415 GGGAGAACATCAAGGAGAAAACCAATACATCAAAAATTGATACCATTTGCTGCCGGATGAAAGC 474

Qy     301 ttcactcaaatggatatctgggggaccgtattcctgaagctcaacactgaagattagaagaagta 360
           |||||||||| ||||||||||| |||| |||||||||| ||||||||||| ||||||
Db     475 TTTACAGAACTTGATCTTGGTGACCGTGTTATGAAACTGAATACAGAGGTCAGAGAATGTA 534

Qy     361 ggtccgtcaacacaggaagggatttttattggcattcaagagtgttggtgcttgttgcc 420
           ||||||||||||| |||||||||||||| ||||||||||||||||||| |||||
Db     535 GGACCTCTAAGCAAAAAAGGGATTTTAATGTGATTACTTTGAGTGTCAGATGAATTCCAAGGCAAT 594

Qy     421 ttggtgtctgtgagagtatacttcaaaaaagtgccattacagtgaagaatctgctatg 480
           ||||| ||||||||| ||||||||||||||||||||| ||||||||| ||
Db     595 CTGGTTTCTGTGCGTGTGTATACTATAAAAAATGCCCTTCGTGGGTACGACACTTGGCTGTC 654

Qy     481 tttccagacacggtacccatgggactccagtccc---tggtggaggttagaggtcttgt 537
           |||| ||| ||  |||||| |||||||||||||||     |||||||||||||||
Db     655 TTTCCCTGACAGCCATCGGAGAAATGCCATGTCAAGGCAGGATATGAGAGAGAAAAATGGCACCTGT 714

Qy     538 gtcaacaattctaaggaggagagatccctccaaggatgtactgcagtacagaaggacggaatgg 597
           |||||||| ||| |||||||||||||||||| |||||| |||||||||||||
Db     715 GTCAACCAATTCTGTGACCGGATGAACCTCCCAAAATGCACCTCCAAAAATGCCACCATTGGCTGCAGATGGGGGAGTGG 774

Qy     598 cttgtacccattgcaatgcttgctctgctatgaaaggaagaaggagggttttatgtgtc 657
           ||||||| ||||| |||||| ||||
Db     775 CTGGTGCCCATCGGAGAAATGCCATGTCAAGGCAGGATATGAGAGAGAAAAATGGCACCTGT 834

Qy     658 caa 660
           |||
Db     835 CAA 837
```

RESULT 10

```
ID     T02949 standard; cDNA; 4529 BP.
AC     T02949;
DT     16-APR-1996 (first entry)
DE     EPH-like receptor protein tyrosine kinase HEK11 cDNA.
KW     human eph-like kinase; therapy; diagnosis; antibody; vector; ss.
OS     Homo sapiens.
FH     Key            Location/Qualifiers
FT     cds            186..3182
FT                    /*tag= a
PN     WO9528484-A1.
PD     26-OCT-1995.
PF     14-APR-1995; U04681.
PR     15-APR-1994; US-229509.
PA     (AMGE-) AMGEN INC.
PI     Fox GM, Jing S, Welcher AA;
DR     WPI; 95-373799/48.
DR     P-PSDB; R85092;
```

CC PCR approach with primers (see T02960-61) based on conserved regions of
CC receptor PTKs and EPH-like receptor PTKs. HEK5, HEK7 and HEK8 show
CC extensive homology to the catalytic domain of chicken EPH-like
CC receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
CC EPH-like receptor. The isolated cDNAs are used for prodn. of
CC recombinant HEKs and chimeric receptors, in hybridisation assays, and
CC to detect abnormalities in HEK receptor genes.

```
SQ     Sequence 4529 BP; 1449 A; 834 C; 1000 G; 1246 T;
```

```
Query Match       45.9%; Score 303; DB 1; Length 4529;
Best Local Similarity 68.0%; Pred. No. 1.1e-90;
Matches  455; Conservative   0; Mismatches 205; Indels   9; Gaps   2;
```

```
Qy      1 tgggaggaagatccagtggtgtggatgaacattacacacccatacaggacttaccagtgtgc 60
          |||||||||||| |||| |||||||||||| |||||||||||| ||||||||||||||
Db    348 TGGGAAGCAAATTAGTGGTTTGGATGAGACAACTATACCCGACGAACATACCAGGTGTGC 407

Qy     61 aatgtcaatggaccacagtgtcaaaacaattggctgagaacacactgggtccccaggaactca 120
          |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db    408 CAAGTCATGGAGACCCCAACCAAAAACAACTGGCCTGCGGGACTACTGGGATTTCCAAAGGCAAT 467

Qy    121 gctcagaagatttatgtggagctcaagttcactctaccacgacactgcaatagcaattccatty 180
          ||||||| || |||||| |||||||||||| ||||||||||||||||||| ||||||
Db    468 GCACAAAGGGATTTTGTTAGAATTGAAATTCACCCTACGGGGATTGTAACAGCTTCCTGGA 527

Qy    181 gtttaggaacttgcaaggagacattcaacctgtactacatggagtctgatgatgatcat 240
          ||||| ||| ||||||| ||||||||||||| ||||||||||||||||||||
Db    528 GTACTGGGACTTGCAGGAAACATTTAATTTGTACTATATGAAACAGACTATGACACT 587

Qy    241 ggggtgaaatttcagagagcatcagtttacaaaggttgaccacctgcagctgatgaaagt 300
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    588 GGCAGGAATAATAAGAACCTCTATGAATGTAAAAAATACAACACCATTGGCTGCAGATGGGGGAGT 647

Qy    301 ttcactcaaatggatatctgggggaccgtattcctgaagctcaacactgaagattagaagaagta 360
          |||||||| ||||| ||| ||||| |||||||||||| ||||||||||||||||||
Db    648 TTTACCCAAGGTCGTGAAAGCAGCTTAAACACTGAGGGTGAGAGGATT 707

Qy    361 ggtccgtcaacacaggaagggatttttattggcattcaagagtgttggtgcttgttgcc 420
          |||||||||||| |||||||||||| ||||||||||||||||||||||| |||||
Db    708 GGACCTTGTGTCCAAAAAGGGATTCTATCCTGCCTTTCAGATGAATGGGGGCCTTGCAATAGCT 767

Qy    421 ttggtgtctgtgagagtatacttcaaaaaagtgccattacagtgaagaatctggctatg 480
          ||||| ||||||||| |||||||||||||||||||||| |||||||||||
Db    768 TTGGTTTCTGTGTCAAAAGTGTACTACAAAAGAAGTGCTCCATTAATTGAGAACTTAGCTATC 827

Qy    481 tttccagacacggtacc---catggactccagtccctgggagttagaggggtcttgt 537
          |||| ||| |||  ||     ||| |||||||||||||||||||||||||||
Db    828 TTTCCAGATACAACAGTGACTGGTTCAGAATTTCCTCTTTAGTCGAGTTGCGAGGGACATGT 887

Qy    538 gtcaacaattctaagga------ggaagatcctccaaggatgtactgcagtacagagaagc 591
          |||||||| |||||||        ||||  ||||| |||| |||||||||||||
Db    888 GTCAGAGAGTGCAGAGAGGAAGAGACGGAGAAAACCCCCAGGAGATGCCGCAGTCCGCAGAAGGA 947

Qy    592 gaatggcttgtacccattggcaatgcttcctgcagctgtgtatgaagaaggaggtttt 651
          |||| ||||| |||||||| ||||
Db    948 GAATGGTTAGTGCCCATTGGAAAAATGTATCTGCAAAGCACAGGCTACGACAGTGTGCAAAAAGGAGAC 1007

Qy    652 atgtgccaa 660
          ||| ||
Db   1008 ACTTGTGAA 1016
```

RESULT 11

```
ID     T32962 standard; cDNA; 2323 BP.
AC     T32962;
DT     11-NOV-1996 (first entry)
DE     Mouse developmental kinase 1 MDK1 T2 clone.
KW     Mouse developmental kinase 1; MDK1 T2; receptor tyrosine kinase;
KW     RTK; signal transduction; probe; diagnosis; gene therapy;
KW     neurodegeneration; neuroproliferation; cancer; ss.
OS     Mus sp.
```

```
FH   Key              Location/Qualifiers
FT   cds              233..2113
FT   signal_peptide   233..316
FT                    /*tag= a
FT                    /*tag= b
FT   mat_peptide      317..2110
FT                    /*tag= c
FT   misc_feature     2017..2113
FT                    /*tag= d
FT                    /*tag= e
FT                    /note="divergent sequence due to alternative
FT                    splicing"
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; US-368776.
PR   03-JAN-1995; US-368776.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
PA   WPI; 96-333988/33.
DR   P-PSDB; W03423.
PS   Example 1; Page 115-116; 128pp; English.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC   coded for truncated versions (W03422 and W03423, respectively) of
CC   the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC   new member of the eck/eph family of receptor tyrosine kinases.
CC   MDK1 T1 and T2 each possess the entire ectodomain, the transmembrane
CC   domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC   but lack the catalytic tyrosine kinase domain. They may have a
CC   modulatory function. The cDNA clones can be used to produce MDK1
CC   T1 and T2, which are useful for screening potential agents for
CC   treatment of diseases characterised by abnormal signal transduction.
SQ   Sequence 2323 BP;    680 A;    497 C;    609 G;    537 T;

Qy   Best Local Similarity   45.2%;
Matches  452;   Conservative   0;   Pred. No. 3.1e-89;   Indels  9;   Gaps  2;

Qy      1  tgggaagagatcaagtggtgtggatgaacattacacaccatcaaggagacttaccaggtgtgc   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    395  TGGGAAGAGAAATTAGTTGGTTTGGATGAGAACTACACTCCGATAAGAGACATACCAGGTGTGC  454

Qy     61  aatgtcatggaccacagtcaaaacacattggctgagaacaaactggttccccaggaactca  120
           ||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Db    455  CAGGTCATGGAGGCCCAACCAGAACAACTGGGCTGCGGACTAACCAGGATTTCTAAAGGCAAC  514

Qy    121  gctcagagagtttatgtgggactccaagttcactcacgagactgcaatagacattccattg  180
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    515  GCACAAAGGAGATTTTGTTGAAGAATTGAAATTCACCTTGAGGGATTGTAATAGTGTTCCCGGA  574

Qy    181  gttttaggaacttgcaaggagagacattcaaactgtactacatggagtctgatgatgatcat  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575  GTCCTGGGACTTGCAAGGAGACGTTTAATTGTACTATTATGAAACAGACTACGACACC  634

Qy    241  ggggtgaaattcgagagcatcaagttacacacaacattgcagctcgatgaaagt  300
           || ||||||||||||||||||||||||||||||||||||| |||||||||| ||
Db    635  GGCAGGGAATATACGAGAAAAACCTTTATGTTAAAATAGAACACCATTGCTGCAGATGAAAGT  694

Qy    301  ttcactcaaatggatctctgggacccgtatctgaagctcaacactcagagattagaagaagta  360
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695  TTCACACAAGGTGACCTTGGTCAAAAGAAAGATGAAGCTGACACACTGAGGGTGAGAGAATT  754

Qy    361  ggtcctgtcaacaagaggaagggattttattttggaacattcaaagatgttggtcttgttgcc  420
           || |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    755  GGACCTTGTCCAAAAAGGGATTCTATTCTCGCCTTTCAGGATGGTAGGGCTTGCATTAGCA  814

Qy    421  ttggtgtctggaggagtacttcaagatggcccatttacaagtgaagaatctggctatg  480
           ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    815  TTGGTTTCTGGTCAAAAGTGTTACTACAGAAGTGCCTGGCACCATTGTTGTTGAGAACTTAGCTGTC  874

Qy    481  tttccagacacggtacc---catggacctcccagtccctggtggaggtggttagagggtcttgt  537
           ||||||||||| |||    ||| ||||||||||| ||||||||||||||||||||||||| |||
Db    875  TTTCCAGATACAGTGACTGCTGGTTCGGAATTTTCTCCTTAGTGCGAGGTCCGTGGGACATGT  934

Qy    538  gtcaacaattctaagga-------ggaagatccctcacaggatgtactgcagtacagaaggc  591
           ||||||||||| |||||        ||||||||||||| ||||||| ||||||||||||||
Db    935  GTCAGCAGTGCCCAGGAAGAGAGCGCAGAAAAAATTCCCCAGAGAATGCATTCCAGTGCAGGA  994

Qy    592  gaatggctgtacccaattggcaagtgttccctgcaaatgctggtcatgaagaaagagttttc  651
           ||||||||||||||||||||||||||||||||| ||||||||||||||||| |||||| |||
Db    995  GAGTGGCGTAGTACCCATTGGAAAATGCATTCGCAAAAGCAGGCGTATCAGCAAAAAGGGGAC  1054

Qy    652  aagtgccaa  660
           ||| ||| ||
Db   1055  ACTTGCGAA  1063
```

```
ID   T32961
T32961
AC   T32961;
DT   11-NOV-1996  (first entry)
DE   Mouse developmental kinase 1 MDK1 T1 clone.
KW   Mouse developmental kinase 1 MDK1 T1; receptor tyrosine kinase;
KW   RTK; signal transduction; probe; diagnosis; gene therapy;
KW   neurodegeneration; neuroproliferation; cancer; ss.
OS   Mus sp.
FH   Key              Location/Qualifiers
FT   cds              233..2065
FT                    /*tag= a
FT   signal_peptide   233..316
FT                    /*tag= b
FT   mat_peptide      317..2062
FT                    /*tag= c
FT   misc_feature     2031..2065
FT                    /*tag= d
FT                    /note="divergent sequence due to alternative
FT                    splicing"
FT   polya_signal     2839..2843
FT                    /*tag= e
PN   WO9621013-A1.
PD   11-JUL-1996.
PF   03-JAN-1996; US-368776.
PR   03-JAN-1995; US-368776.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA   (SUGE-) SUGEN INC.
PI   Clossek T, Millauer B, Ullrich A;
PA   WPI; 96-333988/33.
DR   P-PSDB; W03422.
PS   Example 1; Page 111-12; 128pp; English.
PT   New mouse development kinase 1 gene - used for developing prods. for
PT   diagnosis and treatment of abnormalities in signal transduction
PT   pathways
CC   cDNA cloning using adult mouse brains and Northern blotting
CC   identified 2 clones, MDK1 T1 (T32961) and MDK1 T2 (T32962), that
CC   coded for truncated versions (W03422 and W03423, respectively) of
CC   the novel mouse developmental kinase 1 (MDK1) (see also T32960), a
CC   new member of the eck/eph family of receptor tyrosine kinases.
CC   MDK1 T1 and T2 each possess the entire ectodomain, the transmembrane
CC   domain and part of the juxtamembrane region of MDK1 (see also W03421),
CC   but lack the catalytic tyrosine kinase domain. They may have a
CC   modulatory function. The cDNA clones can be used to produce MDK1
CC   T1 and T2, which are useful for screening potential agents for
CC   treatment of diseases characterised by abnormal signal transduction.
SQ   Sequence 2901 BP;    890 A;    578 C;    650 G;    783 T;

Query Match
Best Local Similarity   45.2%;
Matches  452;   Conservative   0;   Pred. No. 3.4e-89;   Indels  9;   Gaps  2;

Qy      1  tgggaagagatcaagtggtgtggatgaacattacacaccatcaaggagacttaccaggtgtgc   60
```

Db 395 TGGGAAGAAAAATTAGTGGTTTTGGATGAGAAGAACTACACTCCGATAAAGAACATACCAGGTGTGC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 454

Qy   61 aatgtcaatggaccacagtcaaaaacaattggctgagaacaaactgggtccccaggaactca
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 120

Db 455 CAGGTCAATGGAGACCCAACCAACAGAACAACTGGCTGCGGACTAACTGGATTTCTAAAGGCAAC
                                                                                                                                                    514

Qy 121 gctcagaagaagattatgtggagctcaagttcaactctacgagagactgcaatagacattccaatg
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 180

Db 515 GCACAAAAGGATTTTTGTAGAATTGAAAATTCACCTTGAGGGATTGTAATAGTCTTCCCGGA
                                                                                                                                                    574

Qy 181 gttttaggaacttgcaaggagacattcaacctgtactacactggagtctgatgatgatcat
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 240

Db 575 GTCCTGGGAACTTGCAAGGAAACGTTTAATTTGTACTATTATGAAACAGACTACGACACC
                                                                                                                                                    634

Qy 241 gggtgaaatttcgagagcatcagtttacacaagtgcccattaccagtgaagaatctggctatg
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 300

Db 635 GGCAGGAATATACGAGAGAAAACCTTTAATGTTACAAGATAGACACCATTGTCTGCAGATGAAAGT
                                                                                                                                                    694

Qy 301 ttcaactcaaatggatcttggggaccgtatcctgaagctcaacactgagattagaagaagta
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 360

Db 695 TTCACACAAGGTGACCTTGGTGAAAGCTCAACACTGAGGTGAGGGTGAGAGAGATT
                                                                                                                                                    754

Qy 361 ggtcctgtcaacagaagaggatttcatttcaagatgttggtgcttgttgtgcc
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 420

Db 755 GGACCTTTGTCCAAAAAGGGATTCTATCTTCGCCTTTCAGGATGTGAGGGCTTGCCATAGCA
                                                                                                                                                    814

Qy 421 ttggtgtcgtgagagtatacttacacagtgaagaatctgctatg
         |||||||||||||||||||||||||||||||||||||| 480

Db 815 TTGGTTTCGTGTCAAAGTGTACTACAAGAAGTGCTGGACCATTGTTGTGAGAACTTAGCTGTC
                                                                                                                                                    874

Qy 481 ttttccagacacggtacc---catggactccccagtccctggtgggaggttagagggtcttgt
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||| 537

Db 875 TTTCCAGATACAGTGACTGGTTCGGAATTTCGCCTTAGTCGAGGTCCGTGGGACATGT
                                                                                                                                                    934

Qy 538 gtcaacaattctaagga------ggaagatcctccaaggatgtactgcagtacagaaggc
         |||||||||||||||| |||||||||||||||||||||||||||||||||||| 591

Db 935 GTCACAGAGTCCGAGGAAGAGGCAGAAAAATCCCCAGAATGCATTGCAGTGCGCAGAAGGA
                                                                                                                                                    994

Qy 592 gaatggcttgtacccattggcaagtgttcctgctatgaagaaaagaggttt
         |||||||||||||||||||||||||||||||||||||||||||||||||||| 651

Db 995 GAGTGGCTAGTACCCATTGGAAAAATGCATCAGCAAGCAGGCTGAATTGTGAATAGTCTTCCCGGA
                                                                                                                                                    1054

Qy 652 atgtgccaa 660
         ||||||||||
Db 1055 ACTTGCGGA 1063

RESULT 13

FH   Key   Location/Qualifiers
FT   cds   233..3229
FT   signal_peptide   233..316
FT   /*tag= a
FT   317..3226
FT   /*tag= b
FT   mat_peptide   317..3226
FT   /*tag= c
FT   3'utr   3230..3628
FT   /*tag= d
FT   polya_signal   3568..3573
FT   /*tag= e
FT   3'utr   3627..4304
FT   /*tag= f
FT   /note= "alternative 3'-untranslated region,
FT   found in one isolated clone"

CC   A cDNA clone (T32960) codes for mouse developmental kinase 1 (MDK1)
CC   (W03421), a new member of the eck/eph family of receptor tyrosine
CC   kinases (RTKs). To isolate the clone, cDNA from mouse embryos was
CC   subjected to PCR amplification with primers based on conserved
CC   motifs (see also W03426-27) of RTKs. An amplified fragment was used
CC   to screen an 11.5-day-old mouse embryo and an adult mouse brain cDNA
CC   library to obtain the MDK1 clone. 4 Sequence variants (see also
CC   T32961-62 and W03422-25) of MDK1 were also identified. MDK1 nucleic
CC   acids can be used for the recombinant prodn. of MDK1, as probes to
CC   detect MDK1, and for the gene therapy of diseases involving
CC   abnormalities in signal transduction, such as neurodegenerative and
CC   neuroproliferative disorders and cancer.
SQ   Sequence 4304 BP; 1297 A; 875 C; 1043 G; 1089 T;

Query Match   45.2%; Score 298.2; DB 1; Length 4304;
Best Local Similarity   67.6%;
Matches 452; Conservative   0; Mismatches 208; Indels   9; Gaps 2;

Qy   1 tgggaagaagatacagttgttgatgaacattacacaccatcaaggagttaccaggtgtgc   60
         ||||||||||||||||||| |||||| ||||||||||||||||| ||||||||||||||||
Db 395 TGGGAAGAAAAATTAGTGGTTTGGATGAGAAGAACTACACTCCGATAAAGAACATACCAGGTGTGC   454

Qy   61 aatgtcaatggaccacagtcaaaaacaattggctgagaacaaactgggtccccaggaactca   120
         |||||||| |||||| |||| ||||||| ||||| ||||||||||||||||||||||||||||
Db 455 CAGGTCAATGGAGACCCAACCAACAGAACAACTGGCTGCGGACTAACTGGATTTCTAAAGGCAAC   514

Qy 121 gctcagaagaagattatgtggagctcaagttcaactctacgagagactgcaatagacattccaatg   180
         |||||||| ||||||| ||| ||||||||| || || ||||||||||||||| ||||||||||
Db 515 GCACAAAAGGATTTTTGTAGAATTGAAAATTCACCTTGAGGGATTGTAATAGTCTTCCCGGA   574

Qy 181 gttttaggaacttgcaaggagacattcaacctgtactacactggagtctgatgatgatcat   240
         |||||||||||||||||||| || ||||| ||||||||||| ||||| ||||| ||||||||
Db 575 GTCCTGGGAACTTGCAAGGAAACGTTTAATTTGTACTATTATGAAACAGACTACGACACC   634

Qy 241 gggtgaaatttcgagagcatcagtttacacaagtgcccattaccagtgaagaatctggctatg   300
         |||||| ||| || ||||| |||| |||||| || ||||||||||||| ||||| ||||||||
Db 635 GGCAGGAATATACGAGAGAAAACCTTTAATGTTACAAGATAGACACCATTGTCTGCAGATGAAAGT   694

Qy 301 ttcaactcaaatggatcttggggaccgtatcctgaagctcaacactgagattagaagaagta   360
         ||||| ||||| ||| ||||||| |||||||| |||||||||| ||||||| || || ||||
Db 695 TTCACACAAGGTGACCTTGGTGAAAGCTCAACACTGAGGTGAGGGTGAGAGAGATT   754

Qy 361 ggtcctgtcaacagaagaggatttcatttcaagatgttggtgcttgttgtgcc   420
         ||||||||||| |||||| |||||||| |||||||||||||||||| |||||||
Db 755 GGACCTTTGTCCAAAAAGGGATTCTATCTTCGCCTTTCAGGATGTGAGGGCTTGCCATAGCA   814

Qy 421 ttggtgtcgtgagagtatacttacacagtgaagaatctgctatg   480
         |||||| | |||| ||||||||||||||||||||||||| |||||
Db 815 TTGGTTTCGTGTCAAAGTGTACTACAAGAAGTGCTGGACCATTGTTGTGAGAACTTAGCTGTC   874

Qy 481 ttttccagacacggtacc---catggactccccagtccctggtgggaggttagagggtcttgt   537
         |||||| ||| ||||||| ||||||||||||||||||||||||| ||||| ||||||||
Db 875 TTTCCAGATACAGTGACTGGTTCGGAATTTCGCCTTAGTCGAGGTCCGTGGGACATGT   934

Qy 538 gtcaacaattctaagga------ggaagatcctccaaggatgtactgcagtacagaaggc   591
         |||||||||||||||| ||||| |||| |||||||||||||||||| ||||||
Db 935 GTCACAGAGTCCGAGGAAGAGGCAGAAAAATCCCCAGAATGCATTGCAGTGCGCAGAAGGA   994

Qy 592 gaatggcttgtacccattggcaagtgttcctgctatgaagaaaagaggttt   651

```
Db   995 GAGTGGCTAGTACCCATTGGAAAATGCATCTGCAAAGCAGCGCTATCAGCAGCAAAAAGGGGAC 1054
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy   652 atgtgccaa 660
             ||| ||
Db  1055 ACTTCGGAA 1063
```

```
RESULT 14
Q90662
ID   Q90662 standard; cDNA; 3056 BP.
AC   Q90662;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek7' cDNA.
KW   Cek7'; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis; ss.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   cds             2..2134
FT                   /*tag= a
PN   WO9515375-A.
PN   08-JUN-1995.
PD   07-SEP-1994; U10140.
PF   03-DEC-1993; US-162809.
PR   (LJOL-) LA JOLLA CANCER RES FOUND.
PA   Pasquale EB, Sajjadi FG;
PI   WPI; 95-215256/28.
DR   P-PSDB; R75714.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Claim 2; Page 106-109; 129pp; English.
CC   Novel EPH-related PTK cDNA clone Cek7 (given in Q90653) and its
CC   variant Cek7+ (Q90661) were isolated from a chick embryo library in
CC   lambda gt11, and another variant, Cek7' (Q90662), from a chick
CC   embryonic brain cDNA library in lambda gt11. The variants may originate
CC   via alternative splicing of the same gene. Cek7 had the lowest level
CC   of expression of 7 novel Eph-related kinases examined and was barely
CC   detectable in adult tissues.
SQ   Sequence 3056 BP;  871 A;  661 C;  762 G;  762 T;

Query Match             37.2%;  Score 245.2;  DB 1;  Length 3056;
Best Local Similarity   68.2%;
Matches 356;  Conservative  0;  Mismatches 163;  Indels  3;  Gaps  1;

Qy   142 ctcaagttcactctacgagagctgcaatagcacttccattggtttaggaacttgcaaggag 201
             |||||||||||| | ||| |||||| ||||||||||||||||||| |||||||||||||||
Db     2 CTCAAATTCACCCTGAGGGAGCTGTAACAGCCTTCCAGGAGGACTTGGGACTTGCAAGGAG 61

Qy   202 acattcaacctgtactacatggagtctgatgatgatcatggggtgaaattcgagaggcat 261
             |||||||||| |||||||| ||| ||||| |||||||||||||||||||||| || ||||
Db    62 ACTTTTAACATGTACTACTTGAGTCAGATGATGAAGATGGGAGAACATCAGAGAGAAT 121

Qy   262 cagttacaaagattgacacaccattcagctgatgaaagtttcactcaaatggatcttggg 321
             |||| ||||| |||||||||||| |||||| ||| |||||| ||| ||||||||| ||||
Db   122 CAGTACATCAAGATAGATAACCATTGCTGCTGAGAGCTTCACGGAGTTGGACCTTCGGC 181

Qy   322 gaccgtattctgaagctcaacactgagattagagaagtagtcctgtcaacaagaaggga 381
             ||||| ||||||||||||||||||| |||||||||| |||| ||||||||||||||||
Db   182 GACAGAGTTATGAAGTTAAAACACAGAAGTGAGAGATGTTGGGCCCTCTAACAAAAAAGGA 241

Qy   382 ttttattttggcattcaagatgttggtgcttggttgtctgtgagagtatac 441
             ||||| |||||||||| ||||| ||| |||| |||||||||||||| |||||
Db   242 TTTTTACTTGCTTTCACGGATGTCCCAGGACACGGTACCCATG 301

Qy   442 ttcaaaaagtgcccatttacagtgaagaatctggtcgtatgttcccagacacggtacccatg 501
             ||||||||| ||||||||||| |||||| |||| |||||||||| ||||||||||||||||
Db   302 TACAAGAAAATGCCCATCAGTGATCCGCAACCTGGCACGCTTTCCAGATACCATCACAGGA 361

Qy   502 ---gactcccagtccctggtggaggtgttagagggtcttgtgtcaacaattctaaggaggaa 558
                |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||
Db   362 GCAGAATTCTCGCAGCTGCTAGAAGAGTGTCAGGCGTCTGTGTCAACCACTCAGTGACTGAT 421
```

```
Qy   559 gatccctgcaaggatgttcagttcagtacaagagaggccgaatggcttgtactgccaagtgt 618
             | |||| ||||||||||||| | | || ||||||| |||| |||||| | ||||||||||
Db   422 GAGGCACCAAAAGATGCACACTGCAGTTCAGAGGGGACCTGGTGCGGCCCCATTGGGAAGTGT 481

Qy   619 tcctgcaatgctggctatgatgaagaaagagaggtttatgtgccaa 660
             ||||||||||||||| |||||||| ||| | | |||||||||||||
Db   482 TTGTGCAAGGCAGGGTACGAGGAGGAGAACAACACTGCCAA 523
```

```
RESULT 15
Q90653
ID   Q90653 standard; cDNA; 3059 BP.
AC   Q90653;
DT   11-NOV-1995 (first entry)
DE   Eph-related PTK Cek7 cDNA.
KW   Cek7; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis; ss.
OS   Gallus sp.
FH   Key             Location/Qualifiers
FT   cds             2..2170
FT                   /*tag= a
PN   WO9515375-A.
PN   08-JUN-1995.
PD   07-SEP-1994; U10140.
PF   03-DEC-1993; US-162809.
PR   (LJOL-) LA JOLLA CANCER RES FOUND.
PA   Pasquale EB, Sajjadi FG;
PI   WPI; 95-215256/28.
DR   P-PSDB; R75705.
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
PS   Claim 2; Page 44-47; 129pp; English.
CC   Novel EPH-related PTK cDNA clone Cek7 (given in Q90653) and its
CC   variant Cek7+ (Q90661) were isolated from a chick embryo library in
CC   lambda gt11, and another variant, Cek7' (Q90662), from a chick
CC   embryonic brain cDNA library in lambda gt11. The variants may originate
CC   via alternative splicing of the same gene. Cek7 had the lowest level
CC   of expression of 7 novel Eph-related kinases examined and was barely
CC   detectable in adult tissues.
SQ   Sequence 3059 BP;  860 A;  632 C;  758 G;  809 T;

Query Match             37.2%;  Score 245.2;  DB 1;  Length 3059;
Best Local Similarity   68.2%;
Matches 356;  Conservative  0;  Mismatches 163;  Indels  3;  Gaps  1;

Qy   142 ctcaagttcactctacgagagctgcaatagcacttccattggtttaggaacttgcaaggag 201
             |||||||||||| | ||| |||||| ||||||||||||||||||| |||||||||||||||
Db     2 CTCAAATTCACCCTGAGGGAGCTGTAACAGCCTTCCAGGAGGACTTGGGACTTGCAAGGAG 61

Qy   202 acattcaacctgtactacatggagtctgatgatgatcatggggtgaaattcgagaggcat 261
             |||||||||| |||||||| ||| ||||| |||||||||||||||||||||| || ||||
Db    62 ACTTTTAACATGTACTACTTGAGTCAGATGATGAAGATGGGAGAACATCAGAGAGAAT 121

Qy   262 cagttacaaagattgacacaccattcagctgatgaaagtttcactcaaatggatcttggg 321
             |||| ||||| |||||||||||| |||||| ||| |||||| ||| ||||||||| ||||
Db   122 CAGTACATCAAGATAGATAACCATTGCTGCTGAGAGCTTCACGGAGTTGGACCTTCGGC 181

Qy   322 gaccgtattctgaagctcaacactgagattagagaagtagtcctgtcaacaagaaggga 381
             ||||| ||||||||||||||||||| |||||||||| |||| ||||||||||||||||
Db   182 GACAGAGTTATGAAGTTAAAACACAGAAGTGAGAGATGTTGGGCCCTCTAACAAAAAAGGA 241

Qy   382 ttttattttggcattcaagatgttggtgcttggttgtctgtgagagtatac 441
             ||||| |||||||||| ||||| ||| |||| |||||||||||||| |||||
Db   242 TTTTTACTTGCTTTCACGGATGTCCCAGGACACGGTACCCATG 301

Qy   442 ttcaaaaagtgcccatttacagtgaagaatctggtcgtatgttcccagacacggtacccatg 501
             ||||||||| ||||||||||| |||||| |||| |||||||||| ||||||||||||||||
Db   302 TACAAGAAAATGCCCATCAGTGATCCGCAACCTGGCACGCTTTCCAGATACCATCACAGGA 361

Qy   502 ---gactcccagtccctggtggaggtgttagagggtcttgtgtcaacaattctaaggaggaa 558
                |||||||||||||||||||||||||||||| ||||| |||||||||||||||||||
Db   362 GCAGAATTCTCGCAGCTGCTAGAAGAGTGTCAGGCGTCTGTGTCAACCACTCAGTGACTGAT 421
```

```
Qy    559  gatcctcccaaggagtgtactgcagtacagagaaggcgaatggcttgtacccaattggcaaagtgt  618
           || | |||| |||||||| ||| || ||||||| | |||||||||| |||||||
Db    422  GAGGCACCAAAAGAATGCACTGCAGTTCAGAGGGAGAATGGCTGGGTGCCCATTGGGAAGTGT  481

Qy    619  tcctgcaatgctggctatgaagaagaagggtttttatgtgccaa  660
           | |||||||| |||| | || |  |||||
Db    482  TTGTGCAAGGCAGGGTACGAGGAGAGAACAACACCTGCCAA  523
```

Search completed: May 15, 2000, 12:08:49
Job time: 4608 sec

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: May 15, 2000, 11:43:14 ; Search time 1225.64 Seconds (without alignments) -523.843 Million cell updates/sec

Title: US-09-104-340-8
Perfect score: 660
Sequence: 1 tgggaagagatcaagtggtgt........aaagaggttttatgtgcaa 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -48639729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

#	DB	#	DB	#	DB
1:	gb_ba1:*	21:	em_om:*	41:	gb_htg3:*
2:	gb_ba2:*	22:	em_or:*	42:	gb_htg4:*
3:	gb_om:*	23:	em_ov:*	43:	gb_htg5:*
4:	gb_ov:*	24:	em_pat:*	44:	gb_htg6:*
5:	gb_pat:*	25:	em_ph:*	45:	gb_htg7:*
6:	gb_ph:*	26:	em_pl:*	46:	em_htg1:*
7:	gb_pl1:*	27:	em_ro:*	47:	em_htg2:*
8:	gb_pl2:*	28:	em_sts:*	48:	em_htg3:*
9:	gb_pr1:*	29:	em_sy:*	49:	em_hum5:*
10:	gb_pr2:*	30:	em_un:*	50:	gb_pl3:*
11:	gb_pr3:*	31:	em_vi:*	51:	gb_pr5:*
12:	gb_ro:*	32:	gb_htg1:*	52:	gb_htg8:*
13:	gb_sts:*	33:	gb_htg2:*	53:	gb_htg9:*
14:	gb_sy:*	34:	gb_in1:*	54:	gb_htg10:*
15:	gb_un:*	35:	gb_in2:*	55:	gb_htg11:*
16:	gb_vi:*	36:	em_ba1:*	56:	gb_htg12:*
17:	em_fun:*	37:	em_ba2:*	57:	gb_htg13:*
18:	em_hum1:*	38:	em_hum3:*	58:	gb_htg14:*
19:	em_hum2:*	39:	em_hum4:*		
20:	em_in:*	40:	gb_pr4:*		

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	3132	5	A28003	A28003 H.sapiens H
2	660	100.0	3132	5	I68018	I68018 Sequence 9
3	660	100.0	3149	9	HUMHEK	HUMHEK Human recep
4	564	85.5	3077	12	RNU69278	U69278 Rattus norv
5	544.8	82.5	2032	12	MUSMEK4SE	M68515 Mouse eph-r
6	544.8	82.5	3197	12	MUSMEK4	M68513 Mouse eph-r
7	464.8	70.4	3241	4	CHKCEK4	M68514 Chicken eph
8	463.2	70.2	3254	5	I15007	I15007 Sequence 15
9	339.8	51.5	3600	4	CHKCEK8	D38174 Gallus gall
10	330.2	50.0	4242	12	MMSEK	X65138 M.musculus
11	330.2	50.0	4242	12	MSEK	S57168 M.musculus
12	327.6	49.6	3943	12	S57168	S57168 Sek-Eph-rel
13	327.6	49.6	77197	11	MMU58332	U58332 Mus musculu
14	327	49.5	3107	10	U90093	U90093 Human Chrom
15	327	49.5	3348	5	HUMRPTKC	L36645 Homo sapien
16	327	49.5	3592	5	I44522	I44522 Sequence 34
17	318.2	48.2	76022	45	AR062743	AR062743 Sequence
18	317.4	48.1	3531	12	RNEHK1	X78689 R.norvegicu
19	317.4	48.1	3906	5	AR062744	AR062744 Sequence
20	317.4	48.1	4165	5	AR025488	AR025488 Sequence
21	317.4	48.1	4322	5	AR043381	AR043381 Sequence
22	317.4	48.1	4322	12	MMU07357	MMU07357 Mus musculu
23	316	47.9	3042	4	XLSEK1	X91191 X.laevis mR
24	316	47.9	3193	4	XELPAGAAA	L26099 Xenopus lae
25	304.6	46.2	3162	10	HUMRPTKB	L36644 Homo sapien
26	304.6	46.2	3903	10	HSEHK1	X95425 H.sapiens m
27	304.6	46.2	132805	41	AC009425	AC009425 Homo sapi
28	303	45.9	4523	10	HUMRPTK	L36642 Homo sapien
29	303	45.9	116490	32	HSI189K14	AL121966 Homo sapi
30	298.2	45.2	2323	12	MMKIN1T2	X79084 M.musculus
31	298.2	45.2	2901	12	MMKIN1T1	X79083 M.musculus
32	298.2	45.2	4304	4	MMKIN1	X79082 M.musculus
33	297.2	45.0	2640	4	DRAJ5030	AJ005030 Danio rer
34	295	44.7	2006	12	RNU21955	U21955 Rattus norv
35	295	44.7	3208	12	RNU21954	U21954 Rattus norv
36	295	44.7	3943	4	GGY14271	Y14271 Gallus gall
37	291.6	44.2	4124	4	GGCEK7B	U03910 Gallus gall
38	280.8	42.5	4737	12	MM72207	U72207 Mus musculu
39	273	41.4	4577	4	DRU89295	U89295 Danio rerio
40	269.6	40.8	160703	10	HS61A9	AL035703 Human DNA
41	245.2	37.2	3056	5	I15010	I15010 Sequence 21
42	245.2	37.2	3059	5	I15001	I15001 Sequence 3
43	245.2	37.2	3125	5	I15009	I15009 Sequence 19
44	230.2	34.9	63166	55	AC023783	AC023783 Homo sapi
45	219.6	33.3	2820	4	GGCEK8A	Z19059 G.gallus Ce

ALIGNMENTS

```
RESULT 1

LOCUS        A28003        3132 bp    DNA        PAT        25-SEP-1995
DEFINITION   H.saplens HEK gene.
ACCESSION    A28003
VERSION      A28003.1  GI:1247486
KEYWORDS     .
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 3132)
  AUTHORS
  TITLE      A NOVEL RECEPTOR-TYPE TYROSINE KINASE AND USE THEREOF
  JOURNAL    Patent: WO 9300425-A 7 07-JAN-1993;
FEATURES             Location/Qualifiers
     source          1..3132
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            100..3051
                     /gene="HEK"
     CDS             100..3051
                     /gene="HEK"
                     /codon_start=1
                     /protein_id="CAA01906.1"
                     /db_xref="GI:1247487"
                     /db_xref="SWISS-PROT:P29320"
                     /translation="MDCQLSILLLSCSVLDSFGELIPQPSNEVNLLDSKITIQGELGW
                     ISYPSHGWEEISGVDEHYTIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDR
                     ILKLNTEIREVGPVNKGFYLAFQDVGACVALVSVRVYFKCPFTVKNLAFFDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMVCSTEGEWLVPIGKCSCNAGYEERGMCOACRPG
                     FYKALDGNMCAKCPHSSTQEDGSMNCRCENNYFRADKDPPSSMACTRPPSSPRAVIS
                     NINETSVILDWSWPLDITGGRKDVTENIICKKCGWNIKQCEPCSPNVRFLPROFGLNT
                     TVTVTDLLAHTNYTFEIDAVNGVSELSSSPPQFAAVSITTNQAAASPVLIKKDRISR
                     NSISLSWQEPEHPNGIILDYEVKYYEKQEQETSYTILRARGTNVTISSLKPDTIYVLQ
                     IRARTAAGYGTNSRKFEFETSPDSFSISGESSQVVMIAISAAVAILLTVVIIVILGR
                     FCGYKSKHGADEKRLHFGNGHLKLPGLRTVYDPHTYEDPTQAVHEFAKELDATNISID
                     KVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSKPVMIYTEYMENGSLDSFLRKHDAQFTVIOLLGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPAAYTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDISTER
                     TTGDWLNGVRPAHCKEIFTGVEYSSCDTIAKISTDDMKKVGVTVVGPQKKIISSIKAL
                     ETQSKNGPVPV"
BASE COUNT     888 a     709 c     761 g     774 t
ORIGIN
```

```
Query Match     100.0%;  Score 660;  DB 5;  Length 3132;
Best Local Similarity   100.0%;
Matches  660;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 tgggaagagatcagttggtggtggatgaacattacacaccatacaggtgtgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 TGGGAAGAGATCAGTTGGTGTGTGGATGAACATTACACACCATACAGGTGTGC 312

Qy     61 aatgtcatggaccacagtcaaaacaaactgggtgttcccaggaactca 120
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 AATGTCATGGACCACAGTCAAAACAAATTGGGTGTCCCCAGGAACTCA 372

Qy    121 gctcaagaagatttatgtgggagctcaagttcactctcacgagactgcaatagcattccattg 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 GCTCAGAAGAGATTTATGTGGGAGCTCAAGTTCACTCTACGAGACTGCAATAGCATTCCANTG 432

Qy    181 gtttaggaacttgcaaggagaaccattcaacctgtactacatggaagtctggatgatgatcat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 GTTTTAGGAACTTGCAAGGAGAACATTCAAACCTGTACTACATGGAAGTCTGATGATGATCAT 492

Qy    241 ggggtgaaattcgagagcatcagtttacaaagattgacaccattgcagctgatgaaagt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 GGGGTGAAATTCGAGAGCATCAGTTTACAAAGATTGACACCATTGCAGCTGATGAAAGT 552

Qy    301 ttcactcaaatggatcttgggaaccgtattctgaagtccaacactgagattagagaagta 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 TTCACTCAAAATGGATCTTGGGGACCGTATTCTGAAGTCCAAGTCAACACTGAGATTAGAGAAGTA 612

Qy    361 ggtcctgtcaacacagacaagagggtgatttttattttgccatttcaaagtatgttggtgcttgtcttgttgcc 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    613 GGTCCTGTCAACAAGACAGGGGATTTTATTTGCCATTTCAAAGATGTTGGTGCTTGTGTTGTGTGCC 672

Qy    421 tcggtgtctgtgaagatatacttcaaaaagtgcccatttacagtgaagaatctggctatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    673 TTGGGTGTCTGTGAGAGTATACTTCAAAAAGTGCCCATTTACAGTGAAGAATCTGGCTATG 732

Qy    481 tttccagacacggtaccccagtgaactccctggtgaggttagaagggtcttgtgtc 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 TTTCCAGACACGGTACCCCAGTGAACTCCCTGGTGAGGTTAGAGGGGTCTTGTGTC 792

Qy    541 aacaattctaaggaggaagaatcctccaaggatgtactgcagtacagaggaagggcgctt 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 AACAATTCTAAGGAGAAGAATCCTCCAAGGATGTACTGCAGTACACTGCAGTACAGAGGAAGGGCGCTT 852

Qy    601 gtacccattgggaagtgttcctgcaatggtgctatggaagaaagaaaggtttttatgtgcaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    853 GTACCCATTGGGCAAGTGTTCCTGCAATGCTATGGAAGAAAGAGGGTTTTATGTGCCAA 912
```

```
RESULT 2

LOCUS        I68018        3132 bp    DNA        PAT        04-FEB-1998
DEFINITION   Sequence 9 from patent US 5674691.
ACCESSION    I68018
VERSION      I68018.1  GI:2830140
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 3132)
  AUTHORS    Boyd,A.W., Simpson,R.John, Wicks,I., Ward,L.David and Wilkinson,D.
  TITLE      Method of screening for ligands to a receptor-type tyrosine kinase
  JOURNAL    Patent: US 5674691-A 9 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1..3132
                     /organism="unknown"
BASE COUNT     888 a     709 c     761 g     774 t
ORIGIN
```

```
Query Match     100.0%;  Score 660;  DB 5;  Length 3132;
Best Local Similarity   100.0%;
Matches  660;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 tgggaagagatcagttggtggtggatgaacattacacaccatacaggtgtgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 TGGGAAGAGATCAGTTGGTGTGTGGATGAACATTACACACCATACAGGTGTGC 312

Qy     61 aatgtcatggaccacagtcaaaacaaactgggtgttcccaggaactca 120
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 AATGTCATGGACCACAGTCAAAACAAATTGGGTGTCCCCAGGAACTCA 372

Qy    121 gctcaagaagatttatgtgggagctcaagttcactctcacgagactgcaatagcattccattg 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 GCTCAGAAGAGATTTATGTGGGAGCTCAAGTTCACTCTACGAGACTGCAATAGCATTCCANTG 432

Qy    181 gtttaggaacttgcaaggagaaccattcaacctgtactacatggaagtctgatgatgatcat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 GTTTTAGGAACTTGCAAGGAGAACATTCAAACCTGTACTACATGGAAGTCTGATGATGATCAT 492

Qy    241 ggggtgaaattcgagagcatcagtttacaaagattgacaccattgcagctgatgaaagt 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 GGGGTGAAATTCGAGAGCATCAGTTTACAAAGATTGACACCATTGCAGCTGATGAAAGT 552

Qy    301 ttcactcaaatggatcttgggaaccgtattctgaagtccaacactgagattagagaagta 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 ... 360
```

```
RESULT 3
LOCUS       HUMHEK        3149 bp    mRNA            PRI       31-DEC-1994
DEFINITION  Human receptor tyrosine kinase (HEK) mRNA, complete cds.
ACCESSION   M83941
VERSION     M83941.1  GI:183931
KEYWORDS    receptor protein-tyrosine kinase.
SOURCE      Homo sapiens lymphoid tumor cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3149)
  AUTHORS   Wicks,I.P., Wilkinson,D., Salvaris,E. and Boyd,A.W.
  TITLE     Molecular cloning of HEK, the gene encoding a receptor tyrosine
            kinase expressed by human lymphoid tumor cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1611-1615 (1992)
  MEDLINE   92179233
FEATURES             Location/Qualifiers
     source          1..3149
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="LK63"
                     /tissue_type="lymphoid tumor"
     gene            101..3052
                     /gene="HEK"
     CDS             101..3052
                     /gene="HEK"
                     /codon_start=1
                     /product="receptor protein kinase"
                     /protein_id="AAA58633.1"
                     /db_xref="GI:183932"
                     /translation="MDCQLSILLLLSCSVLDSFGELIPQPSNEVNLLDSKTIQGELGW
                     ISPSHGWEEISGVDEHYIPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFT
                     LRDCNSIPLVLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESSTQMLDGDR
                     ILKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPM
                     DSQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQACRPG
                     FYKALDGNMKCAKCPHSSTQEDGSMNCRCENNYFRADKPPSSPRNVIS
                     NINETSVILDWSWPLDTGGRKDVTENIICKKCGWNIKQCEPCSPNVRFLPRQFGLTNT
                     TVTVTDLLAHTNYTFEIDAVNGVSELSSPPRQFAAVSITTNQAAPSPVLTIKKDRTSR
                     NSISLSWQEPEHPNGIILDYEVKYKEKQEDETSYTILRARGTNVTISSSLKPDTIVFQ
                     IRARTAAGYGTNSRKFEFETSPDSFSISESSSQVVMIAISAAVAIILLTVVIVLIGR
                     FCGYKSKHGADEKRLHFGNGHLKLPGLRTYVDPHYEDPNQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGTEKQRDFLGEASIMGQFDHPNI
                     IRLEGVVTKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPPPMDCPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDISTFR
                     TTGDWLNGVRTAHCKEIFTGVEYSSCDTIAKISTDDMKRVGVTVVGPQKIISSIKAL
                     ETQSKNGPVPV"
BASE COUNT      891 a      711 c      768 g      779 t
ORIGIN
```

```
Query Match        100.0%;  Score 660; DB 9; Length 3149;
Best Local Similarity  100.0%;  Pred. No. 4.7e-190;
Matches  660;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      361 ggtcctgtcaacaagaagaggatttattattggcattcaagatgttggtgcttgttgcc 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      613 GGTCCTGTCAACAAGAAGAGGATTTATTATTGGCATTCAAGATGTTGGTGCTTGTTGCC 672

Qy      421 ttggtgtctgtgagagtatactttaaaaagtgcccattttacagtgaagaatctggctatg 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      673 TTGGTGTCTGTGAGAGTATACTTTAAAAAGTGCCCATTTTACAGTGAAGAATCTGGCTATG 732

Qy      481 tttcagacaacggtaccatggaactcccagtccccagttttacagtgaagaaactgggctt 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      733 TTTCAGACAACGGTACCATGGAACTCCCAGTCCCCAGTCCCCAGTTTTACAGTGAAGAAACTGGGCTT 792

Qy      541 aacaaattctaaggaggaggagatcctccaaggatgtactgcagtacgacagaggacattcatg 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      793 AACAAATTCTAAGGAGGAGGAGATCCTCCAAGGATGTACTGCAGTACGACAGAGGACATTCATG 852

Qy      601 gtacccattggcaagtgttcctgcaatgctggctatgtgaaagaagaaagaggtttatgtgccaa 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      853 GTACCCATTGGCAAGTGTTCCTGCAATGCTGGCTATGTGAAAGAAGAAAGAGGTTTATGTGCCAA 912
```

```
RESULT 4
LOCUS       RNU69278        3077 bp    mRNA            ROD       23-MAR-1998
DEFINITION  Rattus norvegicus eph-related receptor tyrosine kinase homolog
            (Rek4) mRNA, complete cds.
ACCESSION   U69278
VERSION     U69278.1  GI:1943913
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McTiernan,C.F. and Feldman,A.M.
  TITLE     IL-1 beta alters the expression of the receptor tyrosine kinase
            gene r-EphA3 in neonatal rat cardiomyocytes
  JOURNAL   Am. J. Physiol. 274 (1), H331-H341 (1998)
  MEDLINE   98120505
REFERENCE   2  (bases 1 to 3077)
  AUTHORS   Li,Y.Y., McTiernan,C.F. and Feldman,A.M.
  TITLE     Direct Submission
```

```
Qy      241 ggggtgaaatttcgagagcatcagtttacaaaagatgaacaccattgaagctaatgaaagt 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      494 GGGGTGAAATTTCGAGAGCATCAGTTTACAAAAGATGAACACCATTGAAGCTAATGAAAGT 553

Qy      301 ttcactcaaatggatctttgggaccgtattctgaagctcaacactgagattagaagagta 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      554 TTCACTCAAAATGGATCTTTGGGACCGTATTCTGAAGCTCAACACTGAGATTAGAGAAGTA 613

Qy      361 ggtcctgtcaacaagaagaggatttattattggcattcaagatgttggtgcttgttgcc 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      614 GGTCCTGTCAACAAGAAGAGGATTTATTATTGGCATTCAAGATGTTGGTGCTTGTTGCC 673

Qy      421 ttggtgtctgtgagagtatactttaaaaagtgcccattttacagtgaagaatctggctatg 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      674 TTGGTGTCTGTGAGAGTATACTTTAAAAAGTGCCCATTTTACAGTGAAGAATCTGGCTATG 733

Qy      481 tttcagacaacggtaccatggaactcccagtccccagttttacagtgaagaaactgggctt 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      734 TTTCAGACAACGGTACCATGGAACTCCCAGTCCCCAGTTTTACAGTGAAGAAACTGGGCTT 793

Qy      541 aacaaattctaaggaggaggagatcctccaaggatgtactgcagtacgacagaggacattcatg 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      794 AACAAATTCTAAGGAGGAGGAGATCCTCCAAGGATGTACTGCAGTACGACAGAGGACATTCATG 853

Qy      601 gtacccattggcaagtgttcctgcaatgctggctatgtgaaagaagaaagaggtttatgtgccaa 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      854 GTACCCATTGGCAAGTGTTCCTGCAATGCTGGCTATGTGAAAGAAGAAAGAGGTTTATGTGCCAA 913
```

JOURNAL   Submitted (01-SEP-1996) Cardiology, University of Pittsburgh, 200
       Lothrop Street, Pittsburgh, PA 15213, USA
COMMENT   On Apr 18, 1997 this sequence version replaced gi:1698721.
FEATURES   Location/Qualifiers
    source   1..3077
                 /organism="Rattus norvegicus"
                 /strain="Sprague-Dawley"
                 /db_xref="taxon:10116"
                 /dev_stage="neonatal"
                 /cell_type="cardiomyocytes"
    gene   1..3077
                 /gene="Rek4"
    CDS   35..2989
                 /gene="Rek4"
                 /codon_start=1
                 /product="eph-related receptor tyrosine kinase homolog"
                 /protein_id="AAC06273.1"
                 /db_xref="GI:1698722"
                 /translation="MDCHLSILILFGGCVLSCSRELSPQPSNEVNLLDSKTIQGELGW
                 ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFT
                 LRDCNSIPLVLGTCKEFFNLYYMESDDHGVKFLEHQFTKIDTIAADESFTQMDLGDR
                 LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPM
                 DSQSLVEVRGSCVNNSKEEDPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPG
                 FYKALDGVKACKCPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVIS
                 NINETSVILDWSWPLDTGGRKDITNIICKKCGWNVRQCEPCSPNVRFLPRQLGLTNT
                 TVTVDLLAHTNYTFEIDAINGVSELSSPPRQFAAVSITTNQAPSPVMTIKKDRTSR
                 NSISLSWQEPEHPNGIILDYEVKYYEKQEOETSYTILRARGTNVTISSLKPDTIYVFQ
                 IRARTAAGYGTNSRKFEFENSPDSFSISGENSHVVMIASAVAIIVTVVTVLGVGR
                 FCGYHKSKHSSDEKRLHFGNGHLRLPGLRTYVDPHTYEDPTQAVHEFAKELDATNIAI
                 DKVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRRDFLGEASIMGQFDHPN
                 IIRLEGVVTKSRPVMIVTEMENGSLDSFLRKHDAQPTVIQLVGMLRGIASGMKYLSD
                 MGYVHRDLAARNILINSNLVCKVYSDFGLSRVLEDDPEAAYTTRGGKIPVRWTSPEATA
                 YRKFTSASDVWSYGIVLWEVMSYGERPYWEMSNQDVIKAVDEGYRLPLPMDCPAALYQ
                 LMLDCWQKDRNNRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLLDQSNVDIATF
                 HTTGDWLNGMRTAHCKELFTGVEYSSCCTIAKISTDDMKKVGVTVVGPQKKIISSIKA
                 LETQSKNGPVPV"
BASE COUNT   877 a   714 c   758 g   728 t
ORIGIN

Query Match   85.5%;   Score 564;   DB 12;   Length 3077;
Best Local Similarity   90.9%;   Pred. No. 7.6e-161;
Matches   600;   Conservative   0;   Mismatches   60;   Indels   0;   Gaps   0;

```
Qy       1 tgggaagaagatcagtgtgtgtgatgaacattacacacccatacattacaccaggtgtgc       60
           ||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db     188 TGGGAAGAAGATCAGTGGTGTGTTGATGAGCATTACACACAATCAGGACTTACCAGGTATGC      247
Qy      61 aatgtcatgaccacagtccaaacaacaatcggctgagaacaacaattggcctcccaggaacctca      120
           || ||||||||||| ||||  ||||||||| |||| || |||||| |||| |||||||||| ||
Db     248 AATGTCATGACCAACAGCACAGCCAAAATAAATTGGCTGAGGACAACAAACTCA          307
Qy     121 gctcagagagatttattgtgggactcaagtccactcctacgagaacgccaatgcagttccattg      180
           ||||  || ||||||||||| ||||||| ||| ||||||| ||||| ||||| ||  ||||||
Db     308 GCTCAGAAGATCATATGTGGAGCTAAAGTTCACACTGCGGGACTGTAAACGCAGCAATTCCATTG      367
Qy     181 gtttaggaacttgtcaaggagcattcaacctgtactacatgtgaagtctgatgatgatcat      240
           ||||| |||||| |||| ||||| ||||| ||| |||||||||||| || ||||||||||||
Db     368 GTTTTGGGGACTTGCAAGGAGCACTTTAACCTGTACTACATGGAGTCTGATGATGATCAC      427
Qy     241 ggggtgaaattcgagaagcatcagttacacaagattgaccacatttgcagctgatgaagt      300
           |||| |||| ||| ||||| ||||| ||| ||||| |||| |||||  |||||||| |||||
Db     428 GGTGTCAAATTCCTAGAGCACCATCAGTTACAAGAAATTGACACATTTGCCGCTGATGAGAAGT      487
Qy     301 ttcactcaaaatggatcttgggaggaccgtattctgaagagagagaagta      360
           |||| ||||| |||||| ||||||||| ||||||||| |||||||| ||||
Db     488 TTCACTCAAAATGGATCTCGGGAGGACCGCATTCTAAAACTCAACATTGAGAGTTAGAGAAGTG      547
Qy     361 gtccctgtccaacaagaaggggaggattttattttgacttttgcttgtgttgtttgcc      420
           || ||| |||||||| ||||| |||||| ||||| |||| ||||||| ||||||||||
Db     548 GGACCAGTCAGTCAATAAAAAGGGGTTTTATTTTGGGCCCTTTCAAGATGTGTTGCCTTGTGTTGCC      607
Qy     421 ttggtgtctgtgaagtatacttcaaaaagtgcccattacagtgaagaatctggctatg      480
           |||||||| ||||| ||| ||||||||||||||||| |||||||||||| ||| || ||||
Db     548 GGACCAGTCAGTCAATAAAAAGGGGTTTTATTTTGGGCCCTTTCAAGATGTGTTGCCTTGTGTTGCC      607
```

RESULT 5
LOCUS   MUSMEK4SE   2032 bp   mRNA   ROD   15-FEB-1994
DEFINITION   Mouse eph-related receptor tyrosine kinase (Mek4) secreted mRNA,
            complete cds.
ACCESSION   M68515
VERSION   M68515.1   GI:454828
KEYWORDS   receptor tyrosine kinase.
SOURCE   Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
            cDNA to mRNA.
  ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutheria;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2032)
AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subraman1,S.
TITLE   Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
JOURNAL   New Biol. 3, 769-778 (1991)
MEDLINE   92031278
COMMENT   On Feb 16, 1994 this sequence version replaced gi:199121.
FEATURES   Location/Qualifiers
    source   1..2032
                 /organism="Mus musculus"
                 /strain="IRC x Swiss Webster"
                 /db_xref="taxon:10090"
                 /dev_stage="11.5 day embryo"
                 /tissue_type="embryo"
    CDS   89..1702
                 /codon_start=1
                 /product="Mek4 secreted"
                 /protein_id="AAA39522.1"
                 /db_xref="GI:454829"
                 /translation="MDCHLSILVLLGCCVLSCSSELSPQPSNEVNLLDSKTIQGELGW
                 ISYPSHGWEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFT
                 LKLNTEIREVGPVNKKGFYLAFQDVGACVALVSVRYFKKCPFTVKNLAMFPDTVPMD
                 LRDCNSIPLVLGTCKETFNLYYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRI
                 SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                 YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                 INETSVILDWSWPLDTGGRKDITFNIICKKCGWNVRQCEPCSPNVRFLPRQLGLINTT
                 VTVDLLAHTNYTFEIDAVNGVSELSSPPRQYAAVSITTNQAPSPVMTIKKDRTSRN
                 SISLSWQEPEHPNGIILDYEVKYYQKQEOETSYTILRARGTNVTISSLKPDTIYVFSF
                 RARTAAGYGTNSRKFEFETSPDCMYFSF"
BASE COUNT   578 a   473 c   460 g   521 t
ORIGIN

Query Match   82.5%;   Score 544.8;   DB 12;   Length 2032;
Best Local Similarity   90.2%;   Pred. No. 5.2e-155;
Matches   595;   Conservative   0;   Mismatches   62;   Indels   3;   Gaps   1;

```
Qy       1 tgggaagaagatcagtggtgtgatgaacattacacacccatacattacacaggtgtgc       60
           ||||||||||||||||||||||||||||| || |||||||||||||||||||||||||||||
Db     242 TGGGAAGAAGATCAGTGGTGTGTTGATGAGCATTACACACAATCAGGACTTACCAGGTGTGC      301
```

polyA_signal   2010..2015
polyA_site   2032

```
RESULT  6

LOCUS       MUSMEK4      3197 bp    mRNA            ROD       15-FEB-1994
DEFINITION  Mouse eph-related receptor tyrosine kinase (Mek4) mRNA, complete
            cds.
ACCESSION   M68513
VERSION     M68513.1  GI:199119
KEYWORDS    .
SOURCE      Mus musculus (strain IRC x Swiss Webster) 11.5 day embryo embryo
            cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3197)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   9201278
FEATURES             Location/Qualifiers
     source          1..3197
                     /organism="Mus musculus"
                     /strain="IRC x Swiss Webster"
                     /db_xref="taxon:10090"
                     /dev_stage="11.5 day embryo"
                     /tissue_type="embryo"
     gene            89..3040
                     /gene="Mek4"
     CDS             89..3040
                     /gene="Mek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA39521.1"
                     /db_xref="GI:199120"
                     /translation="MDCHLSLIVLLGCCVLSCSGGELSPQPSNEVNLLDSKIIQGELGW
                     ISPSHGWEEISGVDEHYIPIRTIQVCNVMDHSQNWLRTNWVPRNSAQKIIVELKFT
                     LRDCNSIPLVLGTCKEIFNLIYMESDDHGVKFREHQETKIDIIAADESFTQMDLGDRI
                     LKLNTIIREVGPVNKRGFYLAFQDVGACVALVSVRYFKKGPETVVNLAMFDTVEMD
                     SQSLVEVRGSCVNNSKEEDPPRMYCSTEGEWLVPIGKCTCNAGYEERGFICQACRPGF
                     YKASDGAAKCAKCPPHSSTQEDGSMNCRCENNYFRAEKDPPSMACARPPSAPRNVISN
                     INETSVILLDWSWPLDTGGRDDITFNIICKKCGGWNVRQCEPCSPNVRFLPRQGLITNTT
                     VTVTDLLAHTINYTFEIDAVNGVSELSSPPRQYAAVSITINQAAPSPVMTIKDRTSRN
                     SISLSWQEPEHPNGIILDYEVKYYQKQEQETSYTILRARGTNWTISSLKPDTTYVFQI
                     RARTAAGYGTNSRKFEEFETSDPSFSISGENSHVVMIAISAAVAIVILTVVIVLVRKF
                     CGYHKSKHSAEEKRLHPGNGHLKLPGLRTVYDPHTVEDPQAVHEFAKELDATNISID
                     KVVGAGEFGEVCSGRLKLPSKKEISVAIKTLKVGYTEKQRDFLGEASIMGQFDHNRI
                     IRLEGVVTKSKPEMIVTEYMENGSLDSFLRKHDAQTVIQLVGMLRGIASGMKYLSDM
                     GYVHRDLAARNILINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAMSY
                     RKFTSASDVWSYGIVLWEVMSGERPYSQMSNQDVIKAVDERYRLPPPMDCPAALYQL
                     MLDCWQKDRNRRPKFEQIVSILDKLIRNPGSLKIITSAAARPSNLLDQSNVDIATFH
                     TTGDWLNGMRFAHCKELFIGVEYSSCDTIAKISTDDMKRVGVTVVGPQKRIISTIKAL
                     ETQSKNGPVPV"

BASE COUNT      907 a    767 c    775 g    748 t
ORIGIN

Query Match     82.5%;   Score 544.8;  DB 12;  Length 3197;
Best Local Similarity   90.2%;
Matches  595;  Conservative   0;  Mismatches  62;  Indels   3;  Gaps   1;
                                   Pred. No. 5.3e-155;
```

```
Qy    1 tgggaaggagatcaagtgtgtgatgaacattacacacccatacaaggacttacacaggtgtgc   60
        ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  242 TGGGAAGGAGATCAGTGTGTTGATGAACATTACACACCCATACAAGGACTTACACAGGTGTGC  301

Qy   61 aatgtcatggaccacagtcaaacaattggctgagaacaacatgggtcccaggaactca      120
        |||||||||||| ||||||| ||||||| ||||||||||||||| |||||||||||||
Db  302 AATGTCATGGATCACAGTCCAAAATAATTAATTGGCTGGTGAGGACAAACTCA             361

Qy  121 gccagaagagaattattgcaagtcaagttcactctacggacgactgcaatagcaattccattg     180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 GCTCAGAAGAGAATTATTGGGAGTCAAGTTCACTCTACGGACGACTGCAATAGCAATTCCATTG  421

Qy  181 gtttaggaaacttgcaagaggaagacattcaaacctgtactacagtctgatgatgatcat      240
        |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 GTTTTTGGGGACTTGCAAGGAGACCTTTAACCTGTACTACAGTCAT   --TGATGATCAT      478

Qy  241 ggggtgaaatttcgagcagcatcaagttcactctacaaagtgacacacccattgcagctgaagt     300
        |||||||||||||||||||| ||||||||||| |||||||||||||||||||||||||||||||
Db  479 GGGGTCAAATTCCGAGCAGCATCAGTTCACAGTTCACGAAGATTGACACCATTGCCGCTGAAAGT  538

Qy  301 ttcactcaaatggatctcttgggaacctgtattctgaagctcaaacacttgaagctcaagaagta   360
        |||| |||||||||| |||||||||| ||||||| ||||||||||||||||||||| |||||||
Db  539 TTCACTCAGATGGATCTCGGGGATCGCATTCTGAAACTCAACACTACTGAGATTAGAGAAGTG     598

Qy  361 ggtcctgtcaacaagaaggaggtttattgggcattcaagatgttggtctgttgcc           420
        ||||||||||||||||||||| ||||||||||||||||| ||||||||||| ||||||
Db  599 GGACCAGTCAACAAAAAGGGGTTTTATTTGGCCTTTCAAGACGTGGGCTGTGTTGCC          658

Qy  421 ttggtgtctgtgaagtatacttcaaaaggtgcccatttacagtgagaaggcgaatggctt      480
        ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  659 TTGGTGTCTGTGAGAGTGTACTTCAAAAAGTGCCCGTTACAGTGAAGAAATCTGGCTAG       718

Qy  481 aacaattctaaggaggaagatcctccagagatgttctcaggcacaacacccagtcccctg      540
        |||||| ||||||||||||||| |||||| |||| |||||| ||||||||||||||||||
Db  719 AATAATTCCAAAGAGGAGGACCCTCCCAGAATGTACTGCTCTACAGAAGGGGAATGGCTG     778

Qy  541 gtacccagtcccctggttggagagtacttaagaagagggttcttgtgtc            600
        ||||| |||||||||||| ||||||| |||||||||||||||||| |||||
Db  779 AATAATTCCAAAGAGGAGGACCCTCCCAGAATGTACTGCTCTACAGAAGGGG           838

Qy  601 gtacccagtcccctggttcctgcaatgctctgctggtcttttatgtgtgccaa        660
        |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  839 GTCCCCATTGGCAAATGCACTTGCAAATGCTGGGTATGAAGAACGAGGTTTCATAATGCCAA     898
```

```
RESULT  7
CHKCEK4
LOCUS       CHKCEK4              3241 bp    mRNA            VRT       15-FEB-1994
DEFINITION  Chicken eph-related receptor tyrosine kinase (Cek4) mRNA, complete
            cds.
ACCESSION   M68514
VERSION     M68514.1  GI:454809
KEYWORDS    receptor tyrosine kinase.
SOURCE      Gallus gallus 10 day embryo embryo cDNA to mRNA.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 3241)
  AUTHORS   Sajjadi,F.G., Pasquale,E.B. and Subramani,S.
  TITLE     Identification of a new eph-related receptor tyrosine kinase gene
            from mouse and chicken that is developmentally regulated and
            encodes at least two forms of the receptor
  JOURNAL   New Biol. 3, 769-778 (1991)
  MEDLINE   92013278
COMMENT     On Feb 16, 1994 this sequence version replaced gi:211446.
FEATURES             Location/Qualifiers
     source          1..3241
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /dev_stage="10 day embryo"
                     /tissue_type="embryo"
     gene            32..2983
                     /gene="Cek4"
     CDS             32..2983
                     /gene="Cek4"
                     /codon_start=1
                     /product="receptor tyrosine kinase"
                     /protein_id="AAA48666.1"
                     /db_xref="GI:211447"
                     /translation="MDRRLPLILLCAALGSAGRLSARPGNEVNLLDSKPIQGELGWI
                     SYPSHGWEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIYVELKFTL
                     RDCNSIPLVLGTCKETFNLYYMESDDHLAKFREHQFTKIDTIAADESFTQMLGDRI
                     LKLNTEVREVGPGVSKKGFYLAFQDVGACVALVSRVYFKKGFCPFTVKNLAMFDTVPMD
                     SQSLVEVRGSCVNHSKEEPKMYCSTEGEWLVPIGKCLCNAGYEERGFDACQACRPGF
                     YKASAGNVKCAKCPHSSTYEDASLNCRCEKNYFRSEKDPPSMACTRPPSAPRNVISN
                     INETSVILLDWSWPIDTGGRDVTNIICKKCGGSSKICEPCSDNVRFLPRQTGLNTT
                     VTVDLLAHYNTFEIDAVNGVSDLSTLSRQFAAVSITNQAAPSPITVIRKDRTSRN
                     SVSLSWQEPEHPNGIILDIYEVKYYEKQEOETSYTILRAKSTNVTISGLKPDTIYVFQI
                     RARTAAARYGTSSRKFEFETSPDSSFSISSENSQVVMIAISAAVAILLTVVVVLIGRF
                     CGYKKSKHGPDEKRLHFGNGHLKLPGLRTVVDPHYTEDPDQNAVHEFAKELDASNISID
                     KVGAGEFGEVCSGRLKLPSKKEISVAIKTLKAGYTEKQRRDFLGEASIMGQFDHPNI
                     IRLEGVTIKSKPVMIVTEYMENGSLDSFLRKHDAQFTVIQLVRMLRQTGLRSN
                     GYHRDLAARNILLINSNLVCKVSDFGLSRVLEDDPEAAYTTRGGKIPIRWTSPEAIAY
                     RKFTSASDAWSYGIVLMEWMSYGERPYWEMSFQDVIKAVDEGYRLPPMDCPAALYQL
                     MLDCWQKDRNNRPKFEQIVSILDKLIRNPSSLKIITNAAARPSNLLLDQSNIDISAFR
                     TAGDWLNGRFRFGQCKGIFTGVEYSSCDTIAKISTDDMKRVGVTVGPQKRIVSSIKTL
                     ETHTKNSPVPV"
     polyA_site      3241
BASE COUNT      916 a    735 c    794 g    796 t
ORIGIN
Qy          1 tgggaagagatcaagtgtgttgggataacacacaccatcaagttaccagtgtgtgc 60
              ||||||||||||| |||||| ||||||||||||||||||||||||||||
Db        182 TGGGAAGAGATTAGTGGTGTTGATGAGACCATTATTACTCCAATCTAGAGAGC 241

Qy         61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactgggtcccaggaactca 120
              ||| ||||| | ||||||||||||||||||| |||| ||||||||| |||||||
Db        242 AATGTTATTGGATCACAGTCAAAAACAATTGGGCTGCGAACAAACTGGATTCCACGCAATTCA 301

Qy        121 gctcagaagaagatttatgtggagctcaagttcactcacgatagacattccattg 180
              ||||||||||||||||||||||||||||||||| |||||| |||| |
Db        302 GCCAGAAGAGATATATGTGGAGCTCAAGTTTACCTTGAGGAGGACTGCAATAGTATCCCTCTA 361
```

```
RESULT  8
I15007
LOCUS       I15007               3254 bp    DNA             PAT       02-APR-1996
DEFINITION  Sequence 15 from patent US 5457048.
ACCESSION   I15007
VERSION     I15007.1  GI:1249915
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3254)
  AUTHORS   Pasquale,E.B. and Sajjadi,F.G.
  TITLE     Eph-related tyrosine kinases, nucleotide sequences and methods of
            use
  JOURNAL   Patent: US 5457048-A 15 10-OCT-1995;
FEATURES             Location/Qualifiers
     source          1..3254
                     /organism="unknown"
BASE COUNT      926 a    737 c    796 g    795 t
ORIGIN
Qy          1 tgggaagagatcaagtgtgttgggatgaacacacaccatcaagttaccagtgtgtgc 60
              ||||||||||||| |||||| ||||||||||||||||||||||||||||
Db        182 TGGGAAGAGATTAGTGGTGTTGATGAGACCATTATTACTCCAATCTAGAGAGC 241

Qy         61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactgggtcccaggaactca 120
              ||| ||||| | ||||||||||||||||||| |||| ||||||||| |||||||
Db        242 AATGTTATTGGATCACAGTCAAAAACAATTGGGCTGCGAACAAACTGGATTCCACGCAATTCA 301

Qy        121 gctcagaagaagatttatgtggagctcaagttcactcacgatagacattccattg 180
              ||||||||||||||||||||||||||||||||| |||||| |||| |
Db        302 GCCAGAAGAGATATATGTGGAGCTCAAGTTTACCTTGAGGAGGACTGCAATAGTATCCCTCTA 361

Qy        181 gtttaggaaacttgcaaggagacacttcaacctgtactacacatggagtctgagatgatgatcat 240
```

Query Match   70.48%; Score 464.8; DB 4; Length 3241;
Best Local Similarity   81.5%;
Matches 538; Conservative   0; Mismatches 122; Indels 0; Gaps 0;
                        Pred. No. 1.2e-130;

Query Match   70.28%; Score 463.2; DB 5; Length 3254;
Best Local Similarity   81.4%;
Matches 537; Conservative   0; Mismatches 123; Indels 0; Gaps 0;
                        Pred. No. 3.6e-130;

RESULT 9

CHKCEK8

```
LOCUS       CHKCEK8      3600 bp     mRNA        VRT        01-FEB-2000
DEFINITION  Gallus gallus mRNA for Cek8, complete cds.
ACCESSION   D38174
VERSION     D38174.1    GI:1236732
KEYWORDS    receptor tyrosine kinase; Cek8.
SOURCE      Gallus gallus spinal cord cDNA to mRNA.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
            Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE   1  (sites)
  AUTHORS   Ohta,K., Nakamura,M., Hirokawa,K., Tanaka,S., Iwama,A., Suda,T.,
            Ando,M. and Tanaka,H.
  TITLE     The receptor tyrosine kinase, Cek8, is transiently expressed on
            subtypes of motoneurons in the spinal cord during development
  JOURNAL   Mech. Dev. 54 (1), 59-69 (1996)
  MEDLINE   96404128
REFERENCE   2  (bases 1 to 3600)
  AUTHORS   Ohta,K.
  JOURNAL   Unpublished (1996)
REFERENCE   3  (bases 1 to 3600)
  AUTHORS   Ohta,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1994) to the DDBJ/EMBL/GenBank databases.
            Kunimasa Ohta, Kumamoto University Graduate School of Medical
            Sciences, Dept. of Neuroscience and Immunology; 4-24-1 Kuhonji,
            Kumamoto, Kumamoto 862, Japan
            (E-mail:ohta9203@gpo.kumamoto-u.ac.jp, Tel:096-344-2111(ex.6754),
            Fax:096-364-3554)
FEATURES             Location/Qualifiers
     source          1..3600
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="spinal cord"
     sig_peptide     3..47
     CDS             3..2963
                     /codon_start=1
                     /product="cek8"
                     /protein_id="BAA07373.1"
                     /db_xref="GI:1236733"
```

```
                     /translation="MAGVPVALLPLLVGVCGAVTSRVYPANEVVLLLDSRSVQGELG
                     WIASPLEGGWEEVSIMDEKNPIRTYQVCNVMEPSQNNWLRDWIPREGAQRVYIEK
                     FTLRDCNSLPGVMGTCKETFNLYYYESNNDKERFIRESQFAKIDTIAADESTQVDIG
                     DRIMKLNTEVRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLIVRNLAQFPDTI
                     TGADTSSLVEVRGSCVNNSEKDVPRKMYCGADGEWLVPIGNCLCNAGYEERNGECQAC
                     KIGYYKALSTDVACAKCPHSTSIWESSTSCTCDRGFFRAENDAASMPCTRPPSAPQN
                     LISNVNETSVNLEWSAPQNKGGRDDISYNVVCKRCAGEPSHCRSGSGGVHFSPQQNG
                     LKTTKVSITDLLAHTNYTFEVWAVNGVSKHNPSQDQAVSVTVTNQAASPITALIQAK
                     EITRHSVALAWLEPDRPNGVILEYEVKYYEKQONERSYRIVRTASRNTDIKILNPLFS
                     YVFHVRARTAAGYGIDGSSGPEFETINVPSPIIGGDSNTLLLVVSVAGSVVLVLVLIAA
                     FVISRRRSKYSKAKQEADEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCIKIE
                     KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKORDFLSEASIMGQFDHPNI
                     IHLEGVVTKCKPVMIITEYMENGSLDAFLRKNDGRETVIQLVGMLRGIGSMKYLSDM
                     SYVHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSGERPYWDMSNQDVIKAIEEGYRLPPPMDCPIALHQL
                     MLDCWQKERSDRPKFGQIVNMLDKLIRNPNSLKRTGSESSRPSTALLLDPSSPEFSAVV
                     SVSDWLQAIKMERYKDNFIAAGYTTLEAVVHMNQDDLARIGITAITHQNKILSSVQAM
                     RSQMQQMHGRMVPV"
     mat_peptide     48..2960
                     /product="Cek8"
BASE COUNT      944 a     876 c     935 g     845 t
ORIGIN
```

Query Match   51.5%; Score 339.8; DB 4; Length 3600;
Best Local Similarity   70.6%; Pred. No. 1.3e-92;
Matches   468; Conservative   0; Mismatches 192; Indels   3; Gaps   1;

```
Qy       1 tgggaagagatcagtgtgtgatgaacattacaacaccatcaggacttacaggtgtgc   60
           ||||| ||  |||| |  | || |      ||        |||| || ||| ||  |||
Db     162 TGGGAGGAAGTGAGCATAATGATGAGAAGAACACTCCGATCCGACACTACCAAGTTTGC  221

Qy      61 aatgtcatgtgaccacagtcaaaacaattggtgagacaaaactgggtcccaggaactca  120
           |||| |  || ||||| ||  || ||   |  | |  ||||||||| |||||||  ||
Db     222 AATGTGATGGAGCCCAGTCCAGTCAGAATAATTGGTTACGAACTGATTGGATTCCCGAGGGG  281

Qy     121 gctcaaagatttatgtgagctcaagttcactctacgagactgcaatagcattccattg  180
            |||| ||| |||| | ||||| ||||| |  |||  || ||    || || |||
Db     282 GCTCAGAGAGGTATATTAATTGAAAAATCAAGTTCACGCTGAGAGACTGCAACAGCCTGCCAGT  341

Qy     181 gttttaggaacttgcaaggagacaatt caaccctgtactacacatggagtctgatgatcat  240
            |||| ||   || || || | ||  | | |||||| ||||| ||| || | ||| || ||
Db     342 GTCATGGGAACTTGCAAAGAGACTTTCAAACCTCTATAACTATGAATCAAACAACGACAAG  401

Qy     241 gggggtgaaatttcgagacgcatcagtttacaaaagattgacaccattgcagcctgatgaaagt  300
           |||| ||||| |  |||| | |||||| |   ||| |||||| ||||| ||||| ||||||
Db     402 GGGGAGTGGCTTGTGCCAGATTGGCAAGTTTGCCTGTGCTGCAGAGAGC  461

Qy     301 ttcactcaaatggatcttgggaccgtattctgaagctcaacactgagattagagagta  360
            || |||||  | | ||  ||||| | |  | || | ||  |  ||| |  || ||| ||
Db     462 TTCACCCAGTGGACATTGGGACCGGATTCATGAAGCTGAATCAGAACTGAGAGGTGGGGACGTG  521

Qy     361 ggtccctgtcaacaagaagggattttattttggttgttcatttcaaagtgttgttgttgcc  420
           ||||| || ||| |||||| ||| |  | |||| |  | ||||||| ||| || |||||
Db     522 GGGCCTCTCAGCAACAAGAAGGGTTTTACTTTGGTTGTTCATTTCAAAAGTGATGAGAGTGCC  581

Qy     421 ttggtgtctgtgagagtatacttccaaaaagtgcccattacagtgaagaatctggctatg  480
           ||| | | | ||||| |||| |||| ||||| ||    |  ||  ||||| | ||||
Db     582 TTGGTGTCTGTTCGTGTCTTCTATAAGAAGTGCCCACTGATAGTTCGAAACCTGGCACAG  641

Qy     481 tttccagacacggtaccccatg---gactccagtccctggtggaggttagaggtcttgt  537
            || ||||| | || ||||||||   ||||||||| |||| | ||| ||| |||   |||
Db     642 TTTCCAGACACCATTACTGGGGCTGATACATCCTCTCTGGTGGAGGTGCGAGGTTCCTGT  701

Qy     538 gtcaacaattctaaggagaggaaagatcctccaaggatgctagtgcagttactggaagagcgaatgg  597
            ||||||| | |  |||| | |   || || || || |||  | |  ||   |  ||||
Db     702 GTCAACAACTCGGACAGGAGAAGGACGTGCCAAAAATGTACTGCGGTGGTGCAGATGGTGAATGG  761

Qy     598 cttgtacccaagtgtcctgcaatgtctggctggccacttactatacagttttatgtgc  657
            |||| ||||||||| ||| |  |||   ||  |  |  ||  |   ||||   || | ||
Db     762 CTGGTACCCATTGGCAACTGTCTGTGCAATGCTGGCTATGAAGAACCAAATGGTGAATGC  821

Qy     658 caa  660
           |||
Db     822 CAA
```

Db   822 CAA 824
          |||

RESULT 10
LOCUS   MMSEK   4242 bp   mRNA   ROD   30-JUN-1993
DEFINITION M.musculus mRNA for tyrosine kinase.
ACCESSION   X65138 S51422
VERSION   X65138.1 GI:54083
KEYWORDS   Sek gene; Tyrosine kinase.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 4242)
AUTHORS   Charnay,P.
TITLE   Direct Submission
JOURNAL   Submitted (26-MAR-1992) P. Charnay, Ecole Normale Superieure, Lab
            de Genetique Moleculaire, URA 1302, 46 Rue d'Ulm, 75230 Paris Cedex
            05, FRANCE
REFERENCE   2 (bases 1 to 4242)
AUTHORS   Gilardi-Hebenstreit,P., Nieto,M.A., Frain,M., Mattei,M.G.,
            Chestier,A., Wilkinson,D.G. and Charnay,P.
TITLE   An Eph-related receptor protein tyrosine kinase gene segmentally
            expressed in the developing mouse hindbrain
JOURNAL   Oncogene 7 (12), 2499-2506 (1992)
MEDLINE   93096484
REMARK   Erratum:[published erratum appears in Oncogene 1993
            Apr;8(4):1103]]
FEATURES   Location/Qualifiers
     source   1..4242
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /dev_stage="8.5 days"
                     /map="band C"
     gene   58..3018
                     /gene="Sek"
     CDS   58..3018
                     /gene="Sek"
                     /note="Eph receptor related"
                     /codon_start=1
                     /product="tyrosine kinase"
                     /protein_id="CAA46268.1"
                     /db_xref="GI:54084"
                     /db_xref="MGD:MGI:98277"
                     /db_xref="SWISS-PROT:Q03137"
                     /translation="MAGIFYFILFSFLFGICDAVTGSRVYPANEVILLDSRSVQGELG
                     WIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRTDWITREGAQRYVIEIK
                     FTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRESQFGKTDLIAADESFTQVDIG
                     DRIMKLNTEIRDVGPLSKKGFYLAFQDVGACLAVLSVRVFYKKCPLTVRNLAQFPDTI
                     TGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEEQNGECQAC
                     KIGYYKALSTDASCAKCPPHSYSVWEGATSCTCDRGFFRADNAASMPCTRPPSAPLN
                     LISNVNETSVNLEWSSPQNTGGRQDIYNVVCKKCGAGDPSKCRPCGSGVHTPQQNG
                     LKTTRVSITDLLAHTNYTFEIWAVNGVSKNFSPDQSVSVTVTTNQAAPSSIALVQAK
                     EVTRYSVALAWLEPDRPNGVILEYEVKYYEKDQONERSYRIVRTAARNTDIKGLNPLTS
                     YVFHVRARTAACGYGDFSEPLEVTNTVVPSRIIGDGANSTVLLVSVSGSVVLVILIAA
                     FVISRRSKYSKAKQEADEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCIKIE
                     IHLEGVVTKCKPVMIIEYMENGSLDAFLRKNDGRFTVIQLVGMLRGIGSGMKYLSDM
                     SYVHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSYGERPYWDMSNQDVIKAIEEGYRLPPPMDCPIALHQL
                     MLDCWQKERSDRPKFGQIVNMLDKLIRNPNSLKRTGSESSRPNTALLDPSSPEFSAVV
                     SVGDWLQAIKMDRYKDNFTAAGYTTLEAVVHMSQDDLARIGITAITHQNKILSSVQAM
                     RTQMQOMHGRMVPV"
     sig_peptide   58..102
                     /gene="Sek"
     mat_peptide   103..3015
                     /gene="Sek"
     misc_feature   2392..2394
                     /product="tyrosine kinase"
                     /gene="Sek"

BASE COUNT   1155 a   1034 c   1052 g   1001 t
ORIGIN   /note="tyrosine autophosphorylated"

RESULT 11

Query Match   50.0%; Score 330.2; DB 12; Length 4242;
Best Local Similarity   69.7%; Pred. No. 1.1e-89;
Matches   462; Conservative   0; Mismatches 198; Indels   3; Gaps   1;

```
Qy     1  tgggaagagaatcagtgtggtgtgatgaaacattaacacaccccatcaggacttaccaggttgtc   60
          ||||||| || ||||| |||||| |||||||| ||||  ||||    || || ||||||||| |||
Db   217  TGGGAGGAAGTGAGCCATTATGGATGAGAAAATACACGATCCGAACCTACCAGGTGTGC        276

Qy    61  aaatctcattggaccacagctcaaaacaaatggcttcggaacaaactgggtccccaggaactca  120
          || |||| ||| ||  |||  || |  ||||| ||| ||||| || ||   |||||| ||| ||
Db   277  AACCTGATGGAAGGCAGCCAGCCAGCACAACTGGCTCGGAACTGACTGGGATCACCCGCGAAGG  336

Qy   121  gctcagaagattatgtggagctcaagttcacttctacgagactgcaatagcattccatg  180
          || || ||| |||||||| ||| ||||| | ||| | ||  || || ||   || |||
Db   337  GCCGCAGAGGGTGTACATTGAAATTAAGTTCACTCTAAGGGACTGCAACAGTCTTCCGGGC  396

Qy   181  gtcttagagaacttgcaaggagacacattcaaccctgtactacatggagtctcgatgatcat  240
          ||  |||| | || || |||| | || |||| || ||| |||| |||  || |||  ||
Db   397  GTCATGGGGACTTGCAAGGAGACGTTAACCTCTACTACTATGAGTCGGACAACGACAAG  456

Qy   241  gggtgaaatttcagaagatcagtttacaaaagattgcacccatttgcagctggtgaaagt  300
          ||  ||| ||| || ||  |||| ||  | ||||| || || || |||||| |||| |||
Db   457  GAGCGCTTCATCAGAGAGAGCCAGTTTGGCAAGACTGACACCATCGCAGCTGACGAGAGC  516

Qy   301  ttcactcaaatggatcctggggaccgtattctgaagctcaacactgagattgaagagta  360
          ||| |  ||||| |||  ||  |||| ||| ||||  || || ||| ||||| || |||
Db   517  TTCACGCAGGTGGACATTGGTGACCGAATCATGAAACTGAACACTGAGATCCGGGATGTA  576

Qy   361  ggtcctgtcaacagaagagggatttttattggcatttcaagatgttggtgttgtgcc  420
          ||  ||| |||| |  ||||||  ||||  ||| ||||| |||||  ||| ||  |||
Db   577  GGGCCACTGAGCAAGAAAGGGTTTTACCTGGCTTTTCAGGATGTGGGGGCTTGCATCGCC  636

Qy   421  ttggtgtctgtgagtatactccaaaaagtgccattacagtgaagaatctggctatg  480
          || ||| || || || |||| || | | || ||||  | || || |||||  | |
Db   637  CTGGTGTCTGTGCGTGTTTTCTACAAGAAGTGTCCACTGACAGTGCGGAACCTGGCCCAG  696

Qy   481  tttccagacacgtaccccatg--gactccagtccctgtgtggaggttagagggttctggt  537
          |||||  |    | || |     |||||  ||  ||| |||| || ||  | || |||
Db   697  TTTCCTGACACCATTACTGGGGCGGACACCTCGTCCCTGGTGGAAGTTCGAGGCTCGTGT  756

Qy   538  gtcaacaaattctaaggaggaagatccctccaaggatgtgactgcagtacagaaggg  597
          || |||| | ||  |  | ||| || ||   |||| || ||||  | |||| ||
Db   757  GTCAACAACTCAGAAGAGAAGGATGTGCCAAAAATGTACTGCGGGGCAGATGGTGAATGG  816

Qy   598  cttgtacccattggcaagtgttcctgcaatgctggctcatgaagaaagaggttttatgtgc  657
          || ||  | ||| ||   | || | || | |||||| | |||||| || ||||| |||
Db   817  CTGGTACCCATTGGCAACTGCCTATGCAACGCTGGGCACGAGGAGCAGAATGGTGAATGC  876

Qy   658  caa  660
          |||
Db   877  CAA  879
```

RESULT 11
LOCUS   S57168   4242 bp   mRNA   ROD   28-JUN-1993
DEFINITION Sek=Eph-related receptor protein tyrosine kinase [mice, mRNA, 4242
            nt].
ACCESSION   S57168
VERSION   S57168.1 GI:298687
KEYWORDS   .
SOURCE   Mus sp.
ORGANISM   Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 4242)
AUTHORS   Gilardi-Hebenstreit,P., Nieto,M.A., Frain,M., Mattei,M.G.,
            Chestier,A., Wilkinson,D.G. and Charnay,P.

TITLE   An Eph-related receptor protein tyrosine kinase gene segmentally
             expressed in the developing mouse hindbrain
JOURNAL   Oncogene 8 (4), 1103 (1993)
MEDLINE   9305393
REMARK   GenBank staff at the National Library of Medicine created this
             entry [NCBI gibbsq 127779] from the original journal article.
COMMENT   This sequence comes from Fig. 5.
             ERRATUM Gilardi-Hebenstreit et al., Oncogene (1992)7, 2499-2507.
FEATURES   Location/Qualifiers
     source   1..4242
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene   58..3018
                     /gene="Sek"
                     /note="Eph-related receptor protein tyrosine kinase"
     CDS   58..3018
                     /gene="Sek"
                     /note="Eph-related receptor protein tyrosine kinase;
                     ERRATUM Gilardi-Hebenstreit et al., Oncogene (1992) 7,
                     2499-2507: This sequence comes from Fig. 5"
                     /codon_start=1
                     /protein_id="AAB25836.1"
                     /db_xref="GI:298688"
                     /translation="MAGIFYFILFSFLFGICDAVTGSRVYPANEVTLLDSRSVQGELG
                     WIASPLEGGWEEVSIMDEKNPIRTYQVCNVMEASQNNWLRTRWITREGAQRVIEIK
                     FTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRESQFGKTDIIAADESFTQVDIG
                     DRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQPDTI
                     TGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEEQNGEQAC
                     KIGIYKALSLDASCAKCPPHSYSWEGATSCTCDRGFFRADNASPCTRPPSAPLN
                     LISNVNETSVNLEWSSPQNTGGRODLSYNVVCKKCGAGDPSKRCPCGSGVHYTPQQNG
                     LKTRVSIDLLAHINYTFEIWAVNGVGKSINPSPDQSVSVTTNQAAPSSIALVOAK
                     EVTRYSVALAWLEPDRPNGVTLEYEVKYYEKDQNERSYRIVRTAARNTDIKGLNPLTS
                     YVFHVRARTAAGYGDFSEEPLEVTTNTVPSRIIGDGANSTVLLVSVSGSSVLIVILIAA
                     FVISRRSRKSKAKQEADEEKHLNQGVRTYVDPFTYEDPNQAVREFAKEIDASCIKIE
                     KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKQRRDFLSEASIMGQFDHPNI
                     IHLEGVTKCKPVMIITEYMENGSLDAFLRKNDGRFTVIQLVGMLRGIGSGMKYLSDM
                     SYVHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSYGERPYWDMSNQDVIKAIEEGYRLPPPMDCPIALHOL
                     MLDCWQKERSDRPKFGQIVNMLDKIRNPNSLKRTGSESSRPNTALLDPSSPEFSAVV
                     SVGDWLQAIKDRIYKDNFTAAGYTTLEAVVHMSQDDLARIGITAITHQNRILSSVQAM
                     RTQMOQMHGRMVPV"
BASE COUNT   1155 a   1034 c   1052 g   1001 t
ORIGIN

Query Match   50.0%; Score 330.2; DB 12; Length 4242;
Best Local Similarity   69.7%; Pred. No. 1.1e-89;
Matches 462; Conservative   0; Mismatches 198; Indels   3; Gaps   1;

Qy   1 tgggaagagatcagttgtgttggatgaacattacacaccccattcaaggagacttacaggtgtgtgcc   60
           ||||| | | |||||||| | |||||||||||||||||||||||||||||||||||||||||||||
Db   217 TGGGGAGGAAGTGAGCAATTATGGATAGAGAAAAATACACCGATCCGAACCTACCAGGTGTGC   276

Qy   61 aatgtcatggaccacagtcaaaacaattggctgagagaacaaactgggtccccaaggaactca   120
           ||| | || | ||| ||| | |||||||||||||||||||| ||||| ||||||||||||||
Db   277 AACGTGATGGAAGCCACAGCAACAACTGGCTGCGGACTGGATCACCCCGGAGGG   336

Qy   121 gctcagagaagattatgtgggactcaagttcactcacgagacatagacattccatg   180
           || ||||| |||||||||| ||| ||||| |||| |||| ||| || |||||||
Db   337 GCGCAGAGGGTGTACATTGAAATTAAGTCACTCTAAGGACTGCAACAGTCTTCCGGC   396

Qy   181 gtttaggaacttgcaaggagggacattccaacctgtactacatggagtctgatgatcat   240
           | ||||||| ||||| |||| ||| ||| ||| ||| |||| ||||| ||||||| ||
Db   397 GTCATCGGGACTTGCAAGGACACGTTAACCTCTACTACTAATGAGTCGGACAACACAAG   456

Qy   241 ggggtgaaattgcagagcatcaagttacaaagagattggacaccaccatggcagctgtgataaagt   300
           | ||| |||||||| || |||||| ||||| ||||||||||| ||| ||||| ||||||||||
Db   457 GAGCGCTTCATCAGAGACGAGAGCCAGTTTGCAAGACTCGACACCATCCCAGCTGTCGATGAGAGC   516

Qy   301 ttcactcaaatgggatcttgggacccgtatctgaagctcaacactgagagataagaagta   360
           |||| || || ||||| |||| | ||||| | |||| ||| ||||| || |||||||
Db   517 TTCACGCAGGTGGACATTGGTGACCGAATCAATGAAACTCAACACTGAGATCCGGGATGTA   576

                                                                              continued

```
RESULT  13
LOCUS       U90093      77197 bp    DNA                    PRI       06-AUG-1997
DEFINITION  Human Chromosome 11 pac pDJ298k13, complete sequence.
ACCESSION   U90093
VERSION     U90093.1  GI:2199513
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 77197)
  AUTHORS   Evans,G.A., Bradbury,P., Brignac,S., Bumelster,R., Davie,J.,
            Aarnilivnsnlnlvckvsdfglsrvleddpeaayttcgkiplrwtapeaiayrkfessas
            Davis,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L.,
            Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
            Hinson,S., Megarity,C., Narayanaswamy,U., Newton,J., O'Brien,K.,
            Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J.,
            Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and
            Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1997) Informatics, Southwestern Medical Center,
            5323 Harry Hines, Dallas, TX 75235-8591, USA
REFERENCE   2  (bases 1 to 77197)
  AUTHORS   Evans,G.A., Bradbury,P., Brignac,S., Bumelster,R., Davie,J.,
            Davis,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L.,
            Garner,H.R., Gordon,M., Gotway,G., Grant,O.,
            Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U.,
            Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S.,
            Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
            Ward,T. and Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1997) Genome Science and Technology Center,
            University of Texas Southwestern Medical Center at Dallas, 5323
            Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE   3  (bases 1 to 77197)
  AUTHORS   Evans,G.A., Bradbury,P., Brignac,S., Bumelster,R., Davie,J.,
            Davis,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L.,
            Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
            Hinson,S., Megarity,C., Narayanaswamy,U., Newton,J., O'Brien,K.,
            Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J.,
            Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and
            Wilson,R.
  TITLE     Template
  JOURNAL   Unpublished
COMMENT     On Jun 19, 1997 this sequence version replaced gi:1890102.
FEATURES             Location/Qualifiers
     source          1..77197
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="pDJ298k13"
                     /chromosome="11"
     repeat_region   complement(5..2483)
                     /rpt_family="L1"
     repeat_region   complement(4324..4378)
                     /rpt_family="Alu"
     repeat_region   complement(6830..7194)
                     /rpt_family="MIR"
     repeat_region   complement(8293..8438)
                     /rpt_family="THE1"
     repeat_region   10594..10816
                     /rpt_family="Alu"
     repeat_region   complement(12868..15221)
                     /rpt_family="Tigger1"
     repeat_region   complement(20435..20710)
                     /rpt_family="Alu"
     repeat_region   complement(21024..25822)
                     /rpt_family="L1"
     repeat_region   complement(25614..25712)
                     /rpt_family="MER25"
     repeat_region   complement(28005..28143)
                     /rpt_family="L1"
     repeat_region   complement(29290..29587)
                     /rpt_family="Alu"
     repeat_region   complement(29648..30102)
                     /rpt_family="L1"
     repeat_region   complement(30104..30528)
                     /rpt_family="Alu"
     repeat_region   complement(35203..35404)
                     /rpt_family="MER25"
     repeat_region   complement(35627..36574)
                     /rpt_family="MER25"
     repeat_region   complement(40418..40703)
                     /rpt_family="LTR13"
```

```
BASE COUNT   1082 a    929 c    957 g    975 t
ORIGIN
```

```
                    MGQFDHPNIIRLEGVVTKRSFPAIGVEAFCPSFLRAGFLNGIQAPHPVTAGGSLPPRI
                    PAGRPVWIVVEEMENGSLDSFLRKHDGHFTVIQOLVGMLRGIIASGMKYLSDMGYVHRDL
                    DAWSYGIVWMEVMSYGERPYWEMSNQDVILSIEEGYRLPAPMCGCPPSLOQLMLHCWQK
                    ERNHRPKFTDIVSFLDKLIRNPSALHTLVEDILLWMPESPGDVPEYPLFVTVGDWLDSI
                    KMGQYKSNFMAAGFTTFDLISRMSIDDIRRIGVILIGHQRRIVSSIQTLRLHMMIOE
                    KGPHV"
```

```
Qy        Best Local Similarity  49.6%;
Matches 460; Conservative  69.5%;  Pred. No. 6.6e-89;  Indels  3;  Gaps  1;
                                   Score 327.6; DB 12;  Length 3943;
Db        0; Mismatches 199;

Qy    1 tgggaagagatcaagtggtgtggatgaacattaccacaccatcaggagttaccaggtgtgc 60
        ||||| |||  || | |  |  ||||| ||||  ||  | |||  ||| |  |||||||||
Db  450 TGGGAYGCCATTACTGAAAATGGATGAACACAATAGGCCCCATACATACATTACCAGGTAYGC 509

Qy   61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactgggtccccaggaactca 120
        ||||||||||| |   |||  ||  ||  |  ||||  ||  ||  ||||||  ||||| |
Db  510 AATGTCATGGACCAACAGAACCAGAACAACTGGCTTCGTACTAACTGGAATCTCGTGAYGCT 569

Qy  121 gctcagaagagattatgtgggagctcagtgttcactctacgagagctgcaatagacattcattg 180
        |||| ||   ||  |  |||  || ||||| |||| |||  || || || |   ||  ||||| |
Db  570 GCACAGAAAATTTATGTGGGAAATGAAGTTCACACTGAGGGATTGTAACAGCATCCCATGG 629

Qy  181 gttttaggaacttgcaaggagacacattcaacctgtactacatggagtctgatgatgatcat 240
        | ||||  ||||||| |  ||  | ||| |  ||| |||| |  ||  |   | | | |||
Db  630 GTCTTGGGACTTGTAAAGAAACCTTTAACCTGTATTATAYATAGAATCCGATGAAATCCCAT 689

Qy  241 ggggtgaaattcgagagcatacagttacacaagagattgcacccattgcagctgatgaaagt 300
        |||  ||   |||  | |  |  | || |||  |||  ||||  |||| |||| || || |||
Db  690 GGGACAAAAATTCAAGCCAAGCCAATATATAAAAGATCGACACAAATTGCTGCTCAATGAGAGT 749

Qy  301 ttcaactcaaatggatcttgggcaccgtattctgaagtcaacactgaaaatcctgagctg 360
        ||  ||  | |||||| |   |||  ||| ||| ||| ||||  ||||| ||| |||||
Db  750 TTTACTCAGATGGATTTGGGTGACCGACATCCTTAAACTCAACACTGAAAATCCCTGAGCTG 809

Qy  361 ggtcctgtcaacagaaggagtatactccatttaattttggccatttcaagagatatgttggtgcttgtgttgcc 420
        ||| ||  |||||||||||  ||| |  | ||||| |||||||||| ||||||| |  || ||| | |||| |||
Db  810 GGGCCTATAGAAAAGGAAAGGAATTTTATTTTTGGCTTTTCAAGAYATGAGCAYGCACGCT 869

Qy  421 ttggtgtctgtgagagtatactccaaaaaagtgcccattcacagtgaagaatctggctatg 480
        |||||| | |||  || |  |  | | || |  ||| ||||| |  || || |||||| |
Db  870 CTGGTCYCAGTCCGAGTTTTCTACAAAAAATGCCCTTCACCGTGAGCTTGGCTATG 929

Qy  481 tttcccagacacggtaccca---tggactcccagtcctggttagaggggtcttgt 537
        |||||||| ||| |||   |   | |||||||||||| |||| ||||| |||||||
Db  930 TTTCCTGATACCATCCCAAAGGGTTGATTCTTCCTCTTTGGTTGTTGAAGTCGGGGGCTCAGT 989

Qy  538 gtcaacaattctaaggaggaagatcctcccaaggatgtgactcgagtgacagaaggcgaatgg 597
        ||||   |   ||  |||  | ||  || | ||| | | | || | ||| ||||| | |||
Db  990 GTAAAAGAGTGCTGAGGAGGCGAGAGATACTCTAAAACTCTACTGTGTGGGAGTGGATTGG 1049

Qy  598 cttgtaccattggcaagtgttcctgcaatgctggctatgtgatgaagaagagagtgtttatgtagtgc 657
        |||||  ||  || ||| | |  || | |||  |||||  | |||||   |||||| |||  |||| | |
Db 1050 CTTGGTTCCTCTTGGAAGGTGTATCTGCAGTACACAGGGTAYGAACAGAAAATCGAGGGGTTCTTGC 1109

Qy  658 ca 659
        ||
Db 1110 CA 1111
```

```
repeat_region   /rpt_family="Alu"
                complement(42659..43816)
repeat_region   /rpt_family="L1"
repeat_region   /rpt_family="L1"
                complement(42946..43193)
repeat_region   /rpt_family="Alu"
                43874..44564
repeat_region   /rpt_family="L1"
                48397..49774
repeat_region   /rpt_family="L1"
                complement(55712..55857)
repeat_region   /rpt_family="L1"
                complement(58953..59284)
repeat_region   /rpt_family="THE1"
                complement(59425..59770)
repeat_region   /rpt_family="THE1"
                63634..63701
repeat_region   /rpt_family="L1"
                64533..64831
repeat_region   /rpt_family="Alu"
                complement(69359..69467)
repeat_region   /rpt_family="MIR"
                complement(75583..75896)
repeat_region   /rpt_family="Alu"
                77142..77197

BASE COUNT    23370 a   13418 c   14087 g   26322 t
ORIGIN
```

```
Query Match           49.6%;   Score 327.6;  DB 11;  Length 77197;
Best Local Similarity 69.5%;   Pred. No. 7.9e-89;
Matches  460;  Conservative  0;  Mismatches  199;  Indels  3;  Gaps  1;

Qy      1 tgggaagaagatcagtggtgtggatgaacattacacaccatcaggaacggtgtgc      60
          ||||| | ||| ||| ||| |||||||| ||         ||  || || || ||||
Db   4625 TGGGATGCCATCACTGAAATGGAIGAACATAAATAGGCCCATTCACACATACCAGGTATGT 4684

Qy     61 aatgtcatggaccacagtcaaaacaacaattggctgagaacaaacaaactgggtcccagg 120
          |||||      ||| |||||||| ||||| |||||||||  ||  |||| |
Db   4685 AATGTAAATGGAACCAAAACAACAACTGCCTTCGTACAAACTGATCTCCGGTGATGCA   4744

Qy    121 gctcagaagagattatgtggagctcaagttcacctctacgagagactgcaatagcaattc 180
          ||||||||  ||||||||||| ||| ||||  ||  |  |  |  |||  ||    ||
Db   4745 GCTCAGAAATTTATGTGCAAAGAAACAATTAAACTCAGTTTTATATGGAATCAGATGAGT  4804

Qy    181 gtttaggaacttgcaaggagacattcaacctgtgactacatcggggtctgatgatgatcat 240
          ||  ||||| ||||||| ||||| ||| | ||| |||||  | ||| |||||||||||
Db   4805 GTCTTGGGGGACTTGCAAAGAAACATTAAAATCACACTAAGGGATTGTAAACACATCCCATG 4864

Qy    241 ggggtgaaatttcgagagcatcagtttacaaaagattgacaaccattgcagctgatgaaagt 300
          ||  | |||||| ||| | ||||| | |||||| || || |||  ||| | |||||||| |||
Db   4865 GGAATTAAATTCAGCACCAAAACCAGTATACAAAGAACGACAAAACAATTGCTCTGATGAGAGT 4924

Qy    301 ttcactcaaatggatcttgggggacgctattctgaagctcaaacactgagattagagagta 360
          || |  ||||||||| | |||||||| ||| |||| ||||||||| | |||| | || ||
Db   4925 TTTACCACAATGGATTGGGTGATCGGCATCCTCAAACTCAAAACTCAACACTGAAAATCGTGAGGTG 4984

Qy    361 ggtcctgtcaacagaagggagttattgggcattcaagatgttgttgcttgttgttgcc 420
          ||  |  | ||| || | ||| ||||  |||||||||| ||||| ||  |
Db   4985 GGGCCTATAGAAGGAAAGGAGATTTTATCTGGCTTTTCAAGACATTGGGGGGCTGCANTGCC 5044

Qy    421 ttggtgtctgtgagagtatacttcaaaaaagtgccccattcaagtgaagaatctggctatg 480
          |||| | |||| || ||| | ||| ||  ||||||||| |||  |  | |||||| | ||
Db   5045 CTGGTTTCAGTCCGTGTTTTCTACAAGAAAATGCCCCTTCACTGTTCGTAACTGGCCATG   5104

Qy    481 tttccagacacggtaccca---tggactcccagtccctggtggaggttagaggtcttgt 537
          |||||  ||||  ||||      |||||||| |||||||||| ||||| ||||| ||||||
Db   5105 TTTCCTGATACCAATTCCACAAGGGTTGATTCCTCCTCTTGGTTGAAGTACGGGGTTCTTGT 5164

Qy    538 gtcaacaattctaagggaggaaagatccctccaaaggatgtgtactgcaataggcgaatgg 597
          || ||  | ||   | |||| ||| | ||  || |  || || || || ||  |  | |||
Db   5165 GTGAAGAGTTCGTGAAGAGCGTGACACTCCTPAAACTGTTATATGTGGAGCTGAITGAGAGATTGG 5224

Qy    598 cttgtaccccattggcaagtgtcctgcaatgctggctggctgatgaagaagaagagtttatgtgc 657
          |||| | |||| ||| ||||||||| |||| |||| ||   ||||| ||||||  |  ||||||
Db   5225 CTGGTTCCTTCTTGGAAGGTGCATTCTGGCCAGTACGAGGAATATGAGGAAATTGAGGGTTCTTGC 5284

Qy    658 ca 659
          ||
Db   5285 CA 5286
```

RESULT 14
HUMRPTKC

```
LOCUS       HUMRPTKC       3107 bp    mRNA            PRI       10-AUG-1995
DEFINITION  Homo sapiens receptor protein-tyrosine kinase (HEK8) mRNA, complete
            cds.
ACCESSION   L36645
VERSION     L36645.1  GI:551613
KEYWORDS    EPH-like receptor PTK; receptor protein-tyrosine kinase.
SOURCE      Homo sapiens (clone library; Stratagene premade library, cat
            #936206) female fetus, 17-18 weeks gestation brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3107)
  AUTHORS   Fox,G.M., Holst,P.L., Chute,H.T., Lindberg,R.A., Janssen,A.M.,
            Basu,R. and Welcher,A.A.
  TITLE     cDNA cloning and tissue distribution of five human EPH-like
            receptor protein-tyrosine kinases
  JOURNAL   Oncogene 10 (5), 897-905 (1995)
  MEDLINE   95206782
FEATURES             Location/Qualifiers
     source          1..3107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="fetus, 17-18 weeks gestation"
                     /sex="female"
                     /clone_lib="Stratagene premade library, cat #936206"
                     /tissue_type="brain"
     mRNA            1..3107
                     /gene="HEK8"
                     /product="receptor protein-tyrosine kinase"
     gene            1..3107
                     /gene="HEK8"
     5'UTR           <1..33
                     /gene="HEK8"
     CDS             34..2994
                     /gene="HEK8"
                     /codon_start=1
                     /product="receptor protein-tyrosine kinase"
                     /protein_id="AAA74246.1"
                     /db_xref="GI:551614"
                     /translation="MAGIFYFALFSCLFGICDAVGSRVYPANEVTLLDSRSVQGELG
                     WIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEPSQONNWLRTDWITREGAQRVYIEIK
                     FTLRDCNSLPFGVMGTCKETFNLIYYESDNDKERFIRENQVKIDTIAADSFTQVDIG
                     DRIMKLNTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTI
                     TGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQAC
                     KIGYYKALSTDATCAKCPHSYSVWEGATSCTCDRGFFRADNAASMPCTRPPSAPLN
                     LISNVNETSVNLEWSSPQNTGGRQDISYNVVCKKCGAGDPSKCRPCGSGVHYTPQONG
                     LKTKVSITDLLAHTNYTFEIWAVNGVSKXNPNPDQSVSVTVTNQQAPSSIALVQAK
                     EVTRYSVALAWLEPDRPNGVILEYEVKYYEKDQNERSYRIVRTAARNTDIKGLRPLTS
                     YVFHVRARTAAGYGDFSEPLEVTTNTVPSRIIGDANSTVLLVSVSGGSVVLVVILAA
                     FVISRRRSKYSKAKQEADEEKHLNQGVRIVDPFTYEDPNQAVREFAKEIDASCIKIE
                     KVIGVGEFGEVCSGRLKVPGKREICVAIKTLKAGYTDKQRRDFLSEASIMGQFDHPNI
                     IHLEGVTKCKPVMIITEYMENGSLDAFLRKNDGRFTVIQLVGMLRGIGSGMKYLSDM
                     SYVHRDLAARNILVNSNLVCKVSDFGMSRVLEDDPEAAYTTRGGKIPIRWTAPEAIAY
                     RKFTSASDVWSYGIVMWEVMSYGERPYWDMSNQDVIKAIEEGYRLPPPMDCPIALHQL
                     MLDCWQKERSDRPKFGQIVNMLDKLIRNPNSLKRTGTESSRPNTALLDPSSPESAVV
                     SVGDWLQAIKMDRYKDNFTAAGYTTLEAVVHVNQEDLARIGITAITHQNKILSSVQAM
                     RTQMQQMHGRMVPV
     3'UTR           2995..3107
                     /gene="HEK8"
BASE COUNT      850 a    720 c    812 g    725 t
ORIGIN
```

```
RESULT   15
LOCUS       I4522
DEFINITION  I4522
            Sequence 34 from patent US 5635177.
ACCESSION   I4522
VERSION     I4522.1  GI:2469235
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3348)
  AUTHORS   Bennett,B.D., Goeddel,D. and Matthews,W.
  TITLE     Protein tyrosine kinase agonist antibodies
  JOURNAL   Patent: US 5635177-A 34 03-JUN-1997;
            Location/Qualifiers
FEATURES
  source   1..3348
            /organism="unknown"
BASE COUNT      924 a    769 c    855 g    800 t
ORIGIN
```

Query Match   49.5%; Score 327; DB 10; Length 3107;
Best Local Similarity   69.4%; Pred. No. 9.9e-89;
Matches 460; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy     1 tgggaagagatcagtggtgtggatgaacatacacaccccatcaggacttaccaggtgtgc     60
         ||||| | ||   || | ||||  |||  ||||||| ||||| |||| ||||||||||||
Db   193 TGGGAGGAAGTGAGTATCATGGATGAATGAAAAATACACCAATCCGAACCTACCAAGTGTGC  252

Qy    61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactggggtcccccaggaactca  120
         |  || ||   ||||| ||||     ||||| |  || |||||  ||| ||||||||||||
Db   253 AATGTGAATGGAACCCAGCCAGAATAACTGGGCTACGAACTGATTGGATCACCCGAGAAGGG  312

Qy   121 gctcagaagatttattgtgagctcaagttcactctacgagactgcaatagcattccattg  180
         | ||| |    ||   | || |||||| |  |  ||| ||| || ||| | |||||||
Db   313 GCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC  372

Qy   181 gttttaggaacttgcaaggagacattcaacctgtactacatggagtctgatgatgatcat  240
         ||| |||| |  ||||| ||||  |||||| |  || ||||| ||  |||||||||||||
Db   373 GTCATGGGGACTTGCAAGGAGACGTTTAACCTGTACTACTATGAATCAGACAACGACAAA  432

Qy   241 ggggtgaaatttcgagagagcatcagtttacacaagaattgacaccattgcaagctgataaagt  300
         |||| || |||  ||||| |||||||  | |||||| || || || ||||| ||| |||||
Db   433 GAGCGTTTCATCAGAGAGAACCAGTTTGTCAAAATTGACACACCATTGCTGCTTGATGAGAGC  492

Qy   301 ttcactcaaatggatcttgggggaccgtattctgaagctcaacactgagagttagagaagta  360
         ||  |||||| ||||| ||||| |||  | |  ||||| |||| || || ||||||| |||
Db   493 TTCACCCAAGTGGACATTGGTGACAGATCATGAAGTGAACACCGAGAGTCCGGGATGTA  552

Qy   361 ggtcctgtcaacaagagagggatttatttggcattcaagatgttggtgcttgtgttgcc  420
         |||||||||||||| ||||| ||||| |  ||||||||||| ||||||| |  ||  |||
Db   553 GGGCCATTAAGCAAAAAAGGGGTTTTACCTGGCTTTTCAGGATGTGCATCGCC  612

Qy   421 ttggtgtctgtgagagtatacttcaaaaagtgcccccatttacagtgaagaatctggcatg  480
         |||||| | |   |||   || | | ||||| |  |  ||||||| |||  |||||| ||
Db   613 CTGGTATCAGTCCGTGTGTTCTATAAAAGTGCCAATCTCACACTCGGCCCAG  672

Qy   481 tttccagacacggtacccatg---gactcccagtccctggttgaggttagagggtcttgt  537
         |||  |||  | |||  ||  |   |||||||| | |  || ||||  |||||| | |||
Db   673 TTTCCTGACACCATCACAGGGGGCCTGATACGTCTTCCCTGGTCGGAAGTTTCGAGGCTCCTGT  732

Qy   538 gtcaacaaattctaaggagggaagatcctccaaggatgtgtacttgactgcagtaggcgaatgg  597
         |||||||  |||| |||| || ||||| | ||||| || ||||| |||| ||| ||||||||
Db   733 GTCAACAACTCAGAAGAGAAGAATGTGCCAAAAATGTACTGTGGGGCAGATGGTGAATGG  792

Qy   598 cttgtacccattggcaagtgttcctgcaatgtctggctatgaagaaagagggttttatgtgc  657
         ||||||||||||||||||||||| |  |||||||| ||| |||| | |||| | | |||||
Db   793 CTGGTACCCATTGGCAACTGCCTATGCAAACGCTGGGGCATGAGGAGCGGGAGGAGAATGC  852

Qy   658 caa  660
         |||
Db   853 CAA  855
```

Query Match   49.5%; Score 327; DB 5; Length 3348;
Best Local Similarity   69.4%; Pred. No. 9.9e-89;
Matches 460; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy     1 tgggaagagatcagtggtgtggatgaacatacacaccccatcaggacttaccaggtgtgc     60
         ||||| | ||   || | ||||  |||  ||||||| ||||| |||| ||||||||||||
Db   160 TGGGAGGAAGTGAGTATCATGGATGAATGAAAAATACACCAATCCGAACCTACCAAGTGTGC  219

Qy    61 aatgtcatggaccacagtcaaaacaattggctgagaacaaactggggtcccccaggaactca  120
         |  || ||   ||||| ||||     ||||| |  || |||||  ||| ||||||||||||
Db   220 AATGTGAATGGAACCCAGCCAGAATAACTGGGCTACGAACTGATTGGATCACCCGAGAAGGG  279

Qy   121 gctcagaagatttattgtgagctcaagttcactctacgagactgcaatagcattccattg  180
         | ||| |    ||   | || |||||| |  |  ||| ||| || ||| | |||||||
Db   280 GCTCAGAGGGTGTATATTGAGATTAAATTCACCTTGAGGGACTGCAATAGTCTTCCGGGC  339

Qy   181 gttttaggaacttgcaaggagacattcaacctgtactacatggagtctgatgatgatcat  240
         ||| |||| |  ||||| ||||  |||||| |  || ||||| ||  |||||||||||||
Db   340 GTCATGGGGACTTGCAAGGAGACGTTTAACCTGTACTACTATGAATCAGACAACGACAAA  399

Qy   241 ggggtgaaatttcgagagagcatcagtttacacaagaattgacaccattgcaagctgataaagt  300
         |||| || |||  ||||| |||||||  | |||||| || || || ||||| ||| |||||
Db   400 GAGCGTTTCATCAGAGAGAACCAGTTTGTCAAAATTGACACACCATTGCTGCTTGATGAGAGC  459

Qy   301 ttcactcaaatggatcttgggggaccgtattctgaagctcaacactgagagttagagaagta  360
         ||  |||||| ||||| ||||| |||  | |  ||||| |||| || || ||||||| |||
Db   460 TTCACCCAAGTGGACATTGGTGACAGATCATGAAGTGAACACCGAGAGTCCGGGATGTA  519

Qy   361 ggtcctgtcaacaagagagggatttatttggcattcaagatgttggtgcttgtgttgcc  420
         |||||||||||||| ||||| ||||| |  ||||||||||| ||||||| |  ||  |||
Db   520 GGGCCATTAAGCAAAAAAGGGGTTTTACCTGGCTTTTCAGGATGTGCATCGCC  579

Qy   421 ttggtgtctgtgagagtatacttcaaaaagtgcccccatttacagtgaagaatctggcatg  480
         |||||| | |   |||   || | | ||||| |  |  ||||||| |||  |||||| ||
Db   580 CTGGTATCAGTCCGTGTGTTCTATAAAAGTGCCAATCTCACACTCGGCCCAG  639

Qy   481 tttccagacacggtacccatg---gactcccagtccctggttgaggttagagggtcttgt  537
         |||  |||  | |||  ||  |   |||||||| | |  || ||||  |||||| | |||
Db   640 TTTCCTGACACCATCACAGGGGGCCTGATACGTCTTCCCTGGTCGGAAGTTTCGAGGCTCCTGT  699

Qy   538 gtcaacaaattctaaggagggaagatcctccaaggatgtgtacttgactgcagtaggcgaatgg  597
         |||||||  |||| |||| || ||||| | ||||| || ||||| |||| ||| ||||||||
Db   700 GTCAACAACTCAGAAGAGAAGAATGTGCCAAAAATGTACTGTGGGGCAGATGGTGAATGG  759

Qy   598 cttgtacccattggcaagtgttcctgcaatgtctggctatgaagaaagagggttttatgtgc  657
         ||||||||||||||||||||||| |  |||||||| ||| |||| | |||| | | |||||
Db   760 CTGGTACCCATTGGCAACTGCCTATGCAAACGCTGGGGCATGAGGAGCGGGAGGAGAATGC  819

Qy   658 caa  660
         |||
Db   820 CAA  822
```

Search completed: May 15, 2000, 11:43:49
Job time: 18656 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2000, 21:16:51 ; Search time 44.2 Seconds
(without alignments)
71.991 Million cell updates/sec

Title: US-09-104-340-1
Perfect score: 1200
Sequence: 1 WEEISGVDEHYTPIRTYQVC.........VPIGKCSCNAGYEERGFMCQ 220
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 14361 seqs, 14463640 residues

Total number of hits satisfying chosen parameters: 143561

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1200	100.0	983	1	US-08-167-919A-10	Sequence 10, Appl
2	1200	100.0	983	2	US-08-449-645A-21	Sequence 21, Appl
3	1200	100.0	983	3	US-08-702-367A-21	Sequence 21, Appl
4	1200	100.0	983	3	US-08-715-106-10	Sequence 10, Appl
5	1200	100.0	983	4	PCT-US95-04681-21	Sequence 21, Appl
6	1136	94.7	983	1	US-08-162-809-16	Sequence 16, Appl
7	1132	94.3	982	1	US-08-673-789-4	Sequence 4, Appl
8	917.5	76.5	986	2	US-08-673-789-3	Sequence 3, Appl
9	914	76.2	948	2	US-08-469-537A-101	Sequence 101, App
10	905.5	75.5	986	2	US-08-449-645A-15	Sequence 15, Appl
11	905.5	75.5	986	2	US-08-702-367A-15	Sequence 15, Appl
12	905.5	75.5	986	4	PCT-US95-04681-15	Sequence 15, Appl
13	905.5	75.5	1104	1	US-08-222-616-36	Sequence 36, Appl
14	905.5	75.5	1104	4	PCT-US95-0428-36	Sequence 36, Appl
15	877	72.2	877	2	US-08-673-789-2	Sequence 2, Appl
16	866.5	72.2	1005	2	US-08-469-537A-103	Sequence 103, App
17	865.5	72.1	928	1	US-08-442-248-2	Sequence 2, Appl
18	865.5	72.1	928	1	US-08-440-815-2	Sequence 2, Appl
19	860.5	71.7	967	2	US-08-449-645A-30	Sequence 30, Appl
20	860.5	71.7	967	2	US-08-702-367A-30	Sequence 30, Appl
21	860.5	71.7	991	2	US-08-449-645A-13	Sequence 13, Appl
22	860.5	71.7	991	2	US-08-702-367A-13	Sequence 13, Appl
23	860.5	71.7	991	3	US-08-449-645A-13	Sequence 13, Appl
24	857.5	71.5	610	2	US-08-469-537A-101	Sequence 101, App
25	857.5	71.5	626	3	US-08-469-537A-101	Sequence 101, App
26	857	71.5	610	4	PCT-US96-00419-5	Sequence 5, Appl
27	856.5	71.4	998	4	PCT-US96-00419-2	Sequence 2, Appl
28	856.5	71.4	998	2	US-08-449-645A-17	Sequence 17, Appl
29	856.5	71.4	998	4	PCT-US95-04681-17	Sequence 17, Appl
30	715	59.6	995	2	US-08-673-789-5	Sequence 5, Appl
31	713	59.4	995	1	US-08-162-809-18	Sequence 18, Appl
32	713	59.4	1011	1	US-08-449-645A-12	Sequence 12, Appl
33	708	59.0	970	1	US-08-162-809-12	Sequence 12, Appl
34	708	59.0	970	2	US-08-449-645A-11	Sequence 11, Appl
35	708	59.0	970	4	PCT-US95-04681-11	Sequence 11, Appl
36	693.5	57.8	984	2	US-08-702-367A-11	Sequence 11, Appl
37	681.5	56.8	710	2	US-08-162-809-6	Sequence 6, Appl
38	681.5	56.8	722	1	US-08-673-789-22	Sequence 22, Appl
39	681.5	56.8	744	1	US-08-162-809-22	Sequence 22, Appl
40	644	53.7	973	1	US-08-162-809-10	Sequence 10, Appl
41	644	53.7	988	1	US-08-449-645A-20	Sequence 20, Appl
42	632	52.7	988	2	US-08-162-809-20	Sequence 20, Appl
43	632	52.7	998	2	US-08-702-367A-20	Sequence 20, Appl
44	632	52.7	998	4	PCT-US95-04681-20	Sequence 20, Appl
45	627	52.2	976	2	US-08-449-645A-18	Sequence 18, Appl

RESULT 1
US-08-167-919A-10
Sequence 10, Application US/08167919A
Patent No. 5674691
GENERAL INFORMATION:
; APPLICANT: Boyd, Andrew W.
; APPLICANT: Simpson, Richard J.
; APPLICANT: Wicks, Ian
; APPLICANT: Ward, Larry D.
; APPLICANT: Wilkinson, David
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,919A
; FILING DATE: 18-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK6841 (AU)
; FILING DATE: 21-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK9992 (AU)
; FILING DATE: 12-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00294
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9159
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

ALIGNMENTS

US-08-167-919A-10

Query Match   100.0%;   Score 1200; DB 1; Length 983;
Best Local Similarity   100.0%;   Pred. No. 3.7e-122;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy    61 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||||||||||||||||||||||||||||||||||||
Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 2
US-08-449-645A-21
; Sequence 21, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-449-645A-21

Query Match   100.0%;   Score 1200; DB 2; Length 983;
Best Local Similarity   100.0%;   Pred. No. 3.7e-122;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy    61 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171
```

RESULT 3
US-08-702-367A-21
; Sequence 21, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 983 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-702-367A-21

Query Match   100.0%;   Score 1200; DB 2; Length 983;
Best Local Similarity   100.0%;   Pred. No. 3.7e-122;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy    61 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy   181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||||||||||||||||||||||||||||||||||||
Db   232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 4
US-08-715-106-10
; Sequence 10, Application US/08715106
; Patent No. 6020306.

Query Match   100.0%;   Score 1200; DB 2; Length 983;
Best Local Similarity   100.0%;   Pred. No. 3.7e-122;
Matches 220; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 111

Qy    61 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 VLGTCKETFNLYMESDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171
```

GENERAL INFORMATION:
    APPLICANT: Boyd, Andrew W.
    APPLICANT: Simpson, Richard J.
    APPLICANT: Wicks, Ian
    APPLICANT: Ward, Larry D.
    APPLICANT: Wilkinson, David
    TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
    TITLE OF INVENTION: AND USE THEREOF
    NUMBER OF SEQUENCES: 14
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Scully, Scott, Murphy & Presser
        STREET: 400 Garden City Plaza
        CITY: Garden City
        STATE: New York
        COUNTRY: U.S.A.
        ZIP: 11530
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/715,106
        FILING DATE: 18-SEP-1996
        CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/167,919
        FILING DATE: 18-APR-1994
        APPLICATION NUMBER: PK6841 (AU)
        FILING DATE: 21-JUN-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PK9992 (AU)
        FILING DATE: 12-DEC-1991
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: PCT/AU92/00294
        FILING DATE: 19-JUN-1992
    ATTORNEY/AGENT INFORMATION:
        NAME: DiGiglio, Frank S.
        REGISTRATION NUMBER: 31,346
        REFERENCE/DOCKET NUMBER: 9159
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 742-4343
        TELEFAX: (516) 742-4366
        TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 10:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 983 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-715-106-10

Query Match
Best Local Similarity   100.0%;   Score 1200; DB 3; Length 983;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy  181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||||||||||||||||||||||||||||||||||||
Db  232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 5
PCT-US95-04681-21
Sequence 21, Application PC/TUS9504681
GENERAL INFORMATION:
    APPLICANT: Fox, Gary M.
    TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
    TITLE OF INVENTION: Kinases
    NUMBER OF SEQUENCES: 28
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Amgen Patent Operations/RBW
        STREET: 1840 Denavilland Drive
        CITY: Thousand Oaks
        STATE: California
        COUNTRY: USA
        ZIP: 91320
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: PCT/US95/04681
        FILING DATE:
        CLASSIFICATION:
    ATTORNEY/AGENT INFORMATION:
        NAME: Winter, Robert B.
        REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 21:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 983 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: protein
PCT-US95-04681-21

Query Match
Best Local Similarity   100.0%;   Score 1200; DB 4; Length 983;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL 111

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  112 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 171

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV 231

Qy  181 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
         ||||||||||||||||||||||||||||||||||||||||
Db  232 NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 271
```

RESULT 6
US-08-162-809-16
Sequence 16, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
    APPLICANT: Pasquale, Elena B.
    APPLICANT: Sajjadi, Fereydoun G.
    TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
    TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
    NUMBER OF SEQUENCES: 26
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: CAMPBELL AND FLORES
        STREET: 4370 La Jolla Village Drive, Suite 700
        CITY: San Diego

STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 983 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-162-809-16

Query Match   94.7%;   Score 1136; DB 1;   Length 983;
Best Local Similarity   94.1%;   Pred. No. 3.5e-115;
Matches 207; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy    1  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNNVVPRNSAQKIIYELKFTLRDCNSIPL  60
         |||||||||||||||||||  ||||||||||||||  ||||||||||| ||||||||||||
Db   51  WEEISGVDEHYTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIIVELKFTLRDCNSIPL  110

Qy   61  VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDIIAADESFTQMDLGDRILKLNTEIREV  120
         ||||||||||||||||||||  ||||||||||||| |||||||||||||||||||| |||
Db  111  VLGTCKETFNLYYMESDDDHLAKFREHQFTKIDIMAADESFTQMDLGDRILKLNTEVREV  170

Qy  121  GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV  180
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  171  GPVSKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV  230

Qy  181  NNSKEEDPPRMYCSTEGEWLVPIGKCLCNAGYEERGFMCQ  220
         | |||| ||:||||||||||||||||| ||||||||||:|
Db  231  NHSKEEEPPKMYCSTEGEWLVPIGKCLCNAGYEERGFACQ  270
```

RESULT 7
US-08-673-789-4
Sequence 4, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,789
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 982
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN

US-08-673-789-4

Query Match   94.3%;   Score 1132; DB 2;   Length 982;
Best Local Similarity   94.1%;   Pred. No. 9.5e-115;
Matches 207; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy    1  WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNNVVPRNSAQKIIYELKFTLRDCNSIPL  60
         |||||||||  |||||||  ||||||||||||||  ||||||||||| ||||||||||||
Db   51  WEEISGVDENTPIRTYQESNVMDHSQNNWLRTNWIPRNSAQKIIVELKFTLRDCNSIPL  110

Qy   61  VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDIIAADESFTQMDLGDRILKLNTEIREV  120
         ||||||||||||||||||||  |||||||||||||||||||||||||||||||||| |||
Db  111  VLGTCKETFNLYYMESDDDHLAKFREHQFTKIDIIAADESFTQMDLGDRILKLNTEVREV  170

Qy  121  GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV  180
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  171  GPVSKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVPMDSQSLVEVRGSCV  230

Qy  181  NNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ  220
         | |||| ||:||||||||||||||||| ||||||||||:|
Db  231  NHSKEEEPPKMYCSTEGEWLVPIGKCLCNAGYEERGFACQ  270
```

RESULT 8
US-08-673-789-3
Sequence 3, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, RENPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,789

```
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/177,812
;     FILING DATE: 04-JAN-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: CAROL M. GRUPPI
;     REGISTRATION NUMBER: 37,341
;     REFERENCE/DOCKET NUMBER: 2026-4105
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) 758-4800
;     TELEFAX: (212) 751-6849
;     TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 986
;     TYPE: AMINO ACID
;     STRANDEDNESS: UNKNOWN
;     TOPOLOGY: UNKNOWN
US-08-673-789-3

Query Match   76.5%;  Score 917.5; DB 2;  Length 986;
Best Local Similarity   74.2%;
Matches  164; Conservative  28; Mismatches  28; Indels   1; Gaps   1;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          |:|:|  ||||:|||||||:||||: |||||||||:|:::||:||||||||||||||:||
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEASQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          |:|||||||||||| ||:|:|  ||||  | ||||||||||||||:||:||||||||||:|
Db  114 VMGTCKETFNLYYYESDNDKERFIRESQFGKTDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDVP-MDSQSLVEVRGSC 179
          ||:.||||||||||||||:||||||||:|||||:|||:||||| |   |.:|||||||||
Db  174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVRGSC 233

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
          ||||:||  |:|||  .|:|||||||.| ||||||||||.||
Db  234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGYEERGFACQ 274

RESULT 9
US-08-469-537A-101
; Sequence 101, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
;   APPLICANT: Maisonpierre, et al.
;   TITLE OF INVENTION: EHK AND ROR TYROSINE
;   TITLE OF INVENTION: KINASES
;   NUMBER OF SEQUENCES: 107
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Regeneron Pharmaceuticals, Inc.
;     STREET: 777 Old Saw Mill River Road
;     CITY: Tarrytown
;     STATE: NY
;     ZIP: 10591
;     COUNTRY: U.S.A.
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/469,537A
;     FILING DATE: 06-JUN-1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: USSN 08/406,247
;     FILING DATE: 17-MAR-1995
;     APPLICATION NUMBER: USSN 08/144,992
;     FILING DATE: 28-OCT-1993
```

```
;     APPLICATION NUMBER: USSN 07/736,559
;     FILING DATE: 26-JUL-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Kempler, Ph.D., Gail M
;     REGISTRATION NUMBER: 32,143
;     REFERENCE/DOCKET NUMBER: REG 070C
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 914-345-7400
;     TELEFAX: 914-345-7721
;     TELEX:
; INFORMATION FOR SEQ ID NO: 101:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 948 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: unknown
;     TOPOLOGY: unknown
;     MOLECULE TYPE: protein
US-08-469-537A-101

Query Match   76.2%;  Score 914.5; DB 2;  Length 948;
Best Local Similarity   74.0%;
Matches  159; Conservative  31; Mismatches  24; Indels   1; Gaps   1;

Qy    1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
          | :|   :|| :|| :||||||||||  : .||||||:|||:  |.||||||||||||||||
Db   56 WDAITEMDEHNRPIHTYQVCNVMEPNQONNWLRTNWISRDAAQKIVEMKFTLRDCNSIPW 115

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
          ||||||||| |||:|||: .|||:|    ||:|:|||||||||||||||||||||||:|||
Db  116 VLGTCKETFTLYYIESDESHGTKFKPSQYIKIDTIAADESFTQMDLGDRILKLNTEVREV 175

Qy  121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDVP-MDSQSLVEVRGSC 179
          ||:    |||||||:|||:||||||||:|||||||:|||||||   :||.:|||||||||
Db  176 GPIERGFYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDIIPRVDSSSLVEVRGSC 235

Qy  180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEE 214
          | |:.|| :|:|||  .|.||:|| |:|.| ||||
Db  236 VKSSEERDTPKLYCGADGDWIVPLGRCICTTGYEE 270

RESULT 10
US-08-449-645A-15
; Sequence 15, Application US/0844964 5A
; Patent No. 5981245
; GENERAL INFORMATION:
;   APPLICANT: Fox, Gary M.
;   TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine
;   TITLE OF INVENTION: Kinases
;   NUMBER OF SEQUENCES: 43
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Amgen Patent Operations/RBW
;     STREET: 1840 Dehavilland Drive
;     CITY: Thousand Oaks
;     STATE: California
;     COUNTRY: USA
;     ZIP: 91320
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/449,645A
;     FILING DATE:
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Winter, Robert B.
;     REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 986 amino acids
;     TYPE: amino acid
```

; MOLECULE TYPE: protein
; TOPOLOGY: linear
US-08-449-645A-15

Query Match   75.5%; Score 905.5; DB 2; Length 986;
Best Local Similarity   73.3%; Pred. No. 4.5e-90;
Matches 162; Conservative 30; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
        |||:| :|| |||||||||||||| :|||||||||:| ::| |:|||||||||| |
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWITREGAQRVVIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
        |:||||||||||| ||:|| | |||:|| ||||||||||||| |:|||| |||||||:|
Db 114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
        ||:.|||||||||||||||:|||||:|||||||:||:|||| :|||| |||:|||||||||
Db 174 GPLSKKGFYLAFQDVGACIALVSVRVYFYKKCPLTVRNLAQFPDTIIGADTSSLVEVRGSC 233

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
        |||| : :| |||| :|||||||| | ||||||| |:||
Db 234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 11
US-08-702-367A-15
Sequence 15, Application US/08702367A
; Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 986 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-367A-15

Query Match   75.5%; Score 905.5; DB 2; Length 986;
Best Local Similarity   73.3%; Pred. No. 4.5e-90;
Matches 162; Conservative 30; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
        |||:| :|| |||||||||||||| :|||||||||:| ::| |:|||||||||| |
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWITREGAQRVVIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
        |:||||||||||| ||:|| | |||:|| ||||||||||||| |:|||| |||||||:|
Db 114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
        ||:.|||||||||||||||:|||||:|||||||:||:|||| :|||| |||:|||||||||
Db 174 GPLSKKGFYLAFQDVGACIALVSVRVYFYKKCPLTVRNLAQFPDTIIGADTSSLVEVRGSC 233

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
        |||| : :| |||| :|||||||| | ||||||| |:||
Db 234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 12
PCT-US95-04681-15
Sequence 15, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 986 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04681-15

Query Match   75.5%; Score 905.5; DB 4; Length 986;
Best Local Similarity   73.3%; Pred. No. 4.5e-90;
Matches 162; Conservative 30; Mismatches 28; Indels 1; Gaps 1;

Qy   1 WEEISGVDEHYTPIRTYQVCNVMDHSQNNWLRTNWVPRNSAQKIVELKFTLRDCNSIPL 60
        |||:| :|| |||||||||||||| :|||||||||:| ::| |:|||||||||| |
Db   54 WEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWITREGAQRVVIEIKFTLRDCNSLPG 113

Qy   61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
        |:||||||||||| ||:|| | |||:|| ||||||||||||| |:|||| |||||||:|
Db 114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy 121 GPVNKKGFYLAFQDVGACVALVSVSRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
        ||:.|||||||||||||||:|||||:|||||||:||:|||| :|||| |||:|||||||||
Db 174 GPLSKKGFYLAFQDVGACIALVSVRVYFYKKCPLTVRNLAQFPDTIIGADTSSLVEVRGSC 233

Qy 180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
        |||| : :| |||| :|||||||| | ||||||| |:||
Db 234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 13
US-08-222-616-36
Sequence 36, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:

Left column

```
APPLICANT:  Bennett, Brian D.
APPLICANT:  Goeddel, David
APPLICANT:  Lee, James M.
APPLICANT:  Matthews, William
APPLICANT:  Tsai, Siao Ping
APPLICANT:  Wood, William I.
TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST
TITLE OF INVENTION:  ANTIBODIES
NUMBER OF SEQUENCES:  42
CORRESPONDENCE ADDRESS:
  ADDRESSEE:  Genentech, Inc.
  STREET:  460 Point San Bruno Blvd
  CITY:  South San Francisco
  STATE:  California
  COUNTRY:  USA
  ZIP:  94080
COMPUTER READABLE FORM:
  MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
  COMPUTER:  IBM PC compatible
  OPERATING SYSTEM:  PC-DOS/MS-DOS
  SOFTWARE:  patin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  US/08/222,616
  FILING DATE:  4-APR-1994
  CLASSIFICATION:  530
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  PCT/US93/00586
  FILING DATE:  22-JAN-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  07/826935
  FILING DATE:  22-JAN-1992
ATTORNEY/AGENT INFORMATION:
  NAME:  Lee, Wendy M.
  REGISTRATION NUMBER:
  REFERENCE/DOCKET NUMBER:  821P2
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  415/225-1994
  TELEFAX:  415/952-9881
  TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO:  36:
SEQUENCE CHARACTERISTICS:
  LENGTH:  1104 amino acids
  TYPE:  amino acid
  TOPOLOGY:  linear
US-08-222-616-36
```

```
Query Match     75.5%;  Score 905.5;  DB 1;  Length 1104;
Best Local Similarity   73.3%;  Pred. No. 5.2e-90;
Matches 162;  Conservative 30;  Mismatches 28;  Indels  1;  Gaps  1;

Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:|| |:|| |||||||||||||||||:|||||||||||||||||||||||||||:|
Db    54 WEEVSIMDEKNTPIRIYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPG 113

Qy    61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |:|||||||||||| |||:| |:|||:|||||||||||:|||||||||||||||||||:|
Db   114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSRVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||||||||||||||||||||||||||||||||||||||||||||| |:||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTIGADTSSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYERGFMCQ 220
         ||||:|:| |:||| ||||||||||||:|||||||||||
Db   234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 14
PCT-US95-04228-36
Sequence 36, Application PCT/US9504228
GENERAL INFORMATION:
  APPLICANT:  Genentech, Inc.
```

Right column

```
APPLICANT:  Bennett, Brian D.
APPLICANT:  Goeddel, David
APPLICANT:  Lee, James M.
APPLICANT:  Matthews, William
APPLICANT:  Tsai, Siao Ping
APPLICANT:  Wood, William I.
TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES:  45
CORRESPONDENCE ADDRESS:
  ADDRESSEE:  Genentech, Inc.
  STREET:  460 Point San Bruno Blvd
  CITY:  South San Francisco
  STATE:  California
  COUNTRY:  USA
  ZIP:  94080
COMPUTER READABLE FORM:
  MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
  COMPUTER:  IBM PC compatible
  OPERATING SYSTEM:  PC-DOS/MS-DOS
  SOFTWARE:  patin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  PCT/US95/04228
  FILING DATE:
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  08/222616
  FILING DATE:  04-APR-1994
ATTORNEY/AGENT INFORMATION:
  NAME:  Wendy M. Lee
  REGISTRATION NUMBER:  00,000
  REFERENCE/DOCKET NUMBER:  821P3PCT
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  415/225-1994
  TELEFAX:  415/952-9881
  TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO:  36:
SEQUENCE CHARACTERISTICS:
  LENGTH:  1104 amino acids
  TYPE:  amino acid
  TOPOLOGY:  linear
PCT-US95-04228-36
```

```
Query Match     75.5%;  Score 905.5;  DB 4;  Length 1104;
Best Local Similarity   73.3%;  Pred. No. 5.2e-90;
Matches 162;  Conservative 30;  Mismatches 28;  Indels  1;  Gaps  1;

Qy     1 WEEISGVDEHYTPIRIYQVCNVMDHSQNNWLRTNWVPRNSAQKIYVELKFTLRDCNSIPL 60
         |||:|| |:|| |||||||||||||||||:|||||||||||||||||||||||||||:|
Db    54 WEEVSIMDEKNTPIRIYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPG 113

Qy    61 VLGTCKETFNLYYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
         |:|||||||||||| |||:| |:|||:|||||||||||:|||||||||||||||||||:|
Db   114 VMGTCKETFNLYYESDNDKERFIRENQFVKIDTIAADESFTQVDIGDRIMKLNTEIRDV 173

Qy   121 GPVNKKGFYLAFQDVGACVALVSRVRYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
         ||||||||||||||||||||||||||||||||||||||||||||| |:||||||||||
Db   174 GPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTIGADTSSLVEVRGSC 233

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYERGFMCQ 220
         ||||:|:| |:||| ||||||||||||:|||||||||||
Db   234 VNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQ 274

RESULT 15
US-08-673-789-2
Sequence 2, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
  APPLICANT:  ZHOU, RENPING; SCHULZ, NICHOLAS,
  APPLICANT:  T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
  APPLICANT:  GEORGE, F.
  TITLE OF INVENTION:  BSK RECEPTOR LIKE
```

~ : TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
~ : TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
~ : TITLE OF INVENTION: METHODS
~ : NUMBER OF SEQUENCES: 14
~ : CORRESPONDENCE ADDRESS:
~ :   ADDRESSEE: MORGAN & FINNEGAN
~ :   STREET: 345 PARK AVENUE
~ :   CITY: NEW YORK
~ :   STATE: NEW YORK
~ :   COUNTRY: USA
~ :   ZIP: 10154
~ : COMPUTER READABLE FORM:
~ :   MEDIUM TYPE: FLOPPY DISK
~ :   COMPUTER: IBM PC COMPATIBLE
~ :   OPERATING SYSTEM: PC-DOS/MS-DOS
~ :   SOFTWARE: WORDPERFECT 5.1
~ : CURRENT APPLICATION DATA:
~ :   APPLICATION NUMBER: US/08/673,789
~ :   FILING DATE:
~ :   CLASSIFICATION: 435
~ : PRIOR APPLICATION DATA:
~ :   APPLICATION NUMBER: 08/177,812
~ :   FILING DATE: 04-JAN-1994
~ : ATTORNEY/AGENT INFORMATION:
~ :   NAME: CAROL M. GRUPPI
~ :   REGISTRATION NUMBER: 37,341
~ :   REFERENCE/DOCKET NUMBER: 2026-4105
~ : TELECOMMUNICATION INFORMATION:
~ :   TELEPHONE: (212) 758-4800
~ :   TELEFAX: (212) 751-6849
~ :   TELEX: 421792
~ : INFORMATION FOR SEQ ID NO: 2:
~ : SEQUENCE CHARACTERISTICS:
~ :   LENGTH: 877
~ :   TYPE: AMINO ACID
~ :   STRANDEDNESS: UNKNOWN
~ :   TOPOLOGY: UNKNOWN
US-08-673-789-2

Query Match   72.2%;   Score 866.5; DB 2;   Length 877;
Best Local Similarity 68.3%; Pred. No. 6.7e-86;
Matches 151; Conservative 35; Mismatches 34;   Indels   1;   Gaps   1;

Qy   1 WEEISGVDEHYTPIRIYQVCNVWDHSQNNWLRTNWVPRNSAQKIYELKFTLRDCNSIPL   60
            |||||| ||| || ||| | ||:|||| |||| :|:||:||||||||||:|||
Db   85 WEEIGEVDENYAPIHTYQVCKVWMEQNQNNWLLTSWISNEGASRIFIELKFTLRDCNSLPG 144

Qy   61 VLGTCKETFNLYIYMESDDDHGVKFREHQFTKIDTIAADESFTQMDLGDRILKLNTEIREV 120
            ||||||||| :|| :|:| ||||||||||||||||||:||||||||||||
Db   145 GLGTCKETFNMYYFESDDENGRSIKENQYIKIDTIAADESFTELDLGDRVMKLNTEVRDV 204

Qy   121 GPVNKKGFYLAFQDVGACVALVSVRVYFKKCPFTVKNLAMFPDTVP-MDSQSLVEVRGSC 179
            ||:||||||||||||||| || ||||||||||||| ||| || :|| ||:|||||| ||
Db   205 GPLSKKGFYLAFQDVGACIALVSVRVYIKKCPSVVRHLAIFPDTITGADSSQLLEVSGSC 264

Qy   180 VNNSKEEDPPRMYCSTEGEWLVPIGKCSCNAGYEERGFMCQ 220
            |:| :||| ||::|| |||||||| |||||||:|||||
Db   265 VNHSVTDDPPKMHCSAEGEWLVPIGKCMCKAGYEEKNGTCQ 305

Search completed: May   9, 2000, 22:25:58
Job time: 4147 sec